2.   The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the sequence shown in Figure 1 (SEQ ID NO:3), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 7 (SEQ ID NO:9), Figure 9 (SEQ ID NO:11), Figure 11 (SEQ ID NO:16), Figure 13 (SEQ ID NO:21), Figure 15 (SEQ ID NO:23), Figure 17 (SEQ ID NO:28), Figure 19 (SEQ ID NO:30), Figure 21 (SEQ ID NO:32), Figure 23 (SEQ ID NO:40), Figure 25 (SEQ ID NO:42), Figure 27 (SEQ ID NO:49), Figure 29 (SEQ ID NO:51), Figure 31 (SEQ ID NO:53), Figure 33 (SEQ ID NO:55), Figure 35 (SEQ ID NO:57), Figure 37 (SEQ ID NO:62), Figure 39 (SEQ ID NO:67), Figure 41 (SEQ ID NO:69), Figure 43 (SEQ ID NO:71), Figure 45 (SEQ ID NO:76), Figure 47 (SEQ ID NO:78), Figure 49 (SEQ ID NO:83), Figure 51 (SEQ ID NO:85), Figure 53 (SEQ ID NO:87), Figure 55 (SEQ ID NO:94), Figure 57 (SEQ ID NO:99), Figure 59 (SEQ ID NO:101), Figure 61 (SEQ ID NO:103), Figure 63 (SEQ ID NO:110), Figure 65 (SEQ ID NO:115), Figure 67 (SEQ ID NO:117), Figure 69 (SEQ ID NO:122), Figure 71 (SEQ ID NO:127), Figure 73 (SEQ ID NO:129), Figure 75 (SEQ ID NO:131), Figure 77 (SEQ ID NO:133), Figure 79 (SEQ ID NO:135), Figure 81 (SEQ ID NO:137), Figure 83 (SEQ ID NO:139), Figure 85 (SEQ ID NO:141), Figure 87 (SEQ ID NO:143), Figure 89 (SEQ ID NO:145), Figure 91 (SEQ ID NO:147), Figure 93 (SEQ ID NO:152), Figure 95 (SEQ ID NO:157), Figure 97 (SEQ ID NO:159), Figure 99 (SEQ ID NO:161), Figure 101 (SEQ ID NO:169), Figure 103 (SEQ ID NO:179), Figure 105 (SEQ ID NO:188), Figure 107 (SEQ ID NO:193), Figure 109 (SEQ ID NO:195), Figure 111 (SEQ ID NO:197), Figure 113 (SEQ ID NO:202), Figure 115 (SEQ ID NO:209), Figure 117 (SEQ ID NO:211), Figure 119 (SEQ ID NO:213), Figure 121 (SEQ ID NO:215), Figure 123 (SEQ ID NO:217), Figure 125 (SEQ ID NO:219), Figure 127 (SEQ ID NO:224), Figure 129 (SEQ ID NO:226), Figure 131 (SEQ ID NO:228), Figure 133 (SEQ ID NO:233), Figure 135 (SEQ ID NO:235), Figure 137 (SEQ ID NO:242), Figure 139 (SEQ ID NO:247), Figure 141 (SEQ ID NO:252), Figure 143 (SEQ ID NO:259), Figure 145 (SEQ ID NO:264), Figure 147 (SEQ ID NO:266), Figure 149 (SEQ ID NO:268), Figure 151 (SEQ ID NO:270), Figure 153 (SEQ ID NO:272), Figure 155 (SEQ ID NO:274), Figure 157 (SEQ ID NO:276), Figure 159 (SEQ ID NO:281), Figure 161 (SEQ ID NO:286), Figure 163 (SEQ ID NO:291), Figure 165 (SEQ ID NO:296), Figure 167 (SEQ ID NO:301), Figure 169 (SEQ ID NO:303), Figure 171 (SEQ ID NO:305), Figure 173 (SEQ ID NO:307), Figure 175 (SEQ ID NO:309), Figure 177 (SEQ ID NO:314), Figure 179 (SEQ ID NO:316), Figure 181 (SEQ ID NO:321), Figure 183 (SEQ ID NO:323), Figure 185 (SEQ ID NO:325), Figure 187 (SEQ ID NO:327), Figure 189 (SEQ ID NO:329), Figure 191 (SEQ ID NO:331), Figure 193 (SEQ ID NO:333), Figure 195 (SEQ ID NO:335), Figure 197 (SEQ ID NO:337), Figure 199 (SEQ ID NO:339), Figure 201 (SEQ ID NO:346), Figure 203 (SEQ ID NO:351), Figure 205 (SEQ ID NO:353), Figure 207 (SEQ ID NO:355), Figure 209 (SEQ ID NO:357), Figure 211 (SEQ ID NO:363), Figure 213 (SEQ ID NO:365), Figure 215 (SEQ ID NO:371), Figure 217 (SEQ ID NO:373), Figure 219 (SEQ ID NO:375), Figure 221 (SEQ ID NO:377), Figure 223 (SEQ ID NO:382), Figure 225 (SEQ ID NO:384), Figure 227 (SEQ ID NO:389), Figure 229 (SEQ ID NO:394), Figure 231 (SEQ ID NO:396), Figure 233 (SEQ ID NO:401), Figure 235 (SEQ ID NO:405), Figure 237 (SEQ ID NO:409), Figure 239 (SEQ ID NO:414), Figure 241 (SEQ ID NO:422), Figure 242 (SEQ ID NO:428) and Figure 245 (SEQ ID NO:430).

3.   The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the full-length coding sequence of the sequence shown in Figure 1 (SEQ ID NO:3), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 7 (SEQ ID NO:9), Figure 9 (SEQ ID NO:11), Figure 11 (SEQ ID NO:16), Figure 13 (SEQ ID NO:21), Figure 15 (SEQ ID NO:23), Figure 17 (SEQ ID NO:28), Figure 19 (SEQ ID NO:30), Figure 21 (SEQ ID NO:32), Figure 23 (SEQ ID NO:40), Figure 25 (SEQ ID NO:42), Figure 27 (SEQ ID NO:49), Figure 29 (SEQ ID NO:51), Figure 31 (SEQ ID NO:53), Figure 33 (SEQ ID NO:55), Figure 35 (SEQ ID NO:57), Figure 37 (SEQ ID NO:62), Figure 39 (SEQ ID NO:67), Figure 41 (SEQ ID NO:69), Figure 43 (SEQ ID NO:71), Figure 45 (SEQ ID NO:76), Figure 47 (SEQ ID NO:78), Figure 49 (SEQ ID NO:83), Figure 51 (SEQ ID NO:85), Figure 53 (SEQ ID NO:87), Figure 55 (SEQ ID NO:94), Figure 57 (SEQ ID NO:99), Figure 59 (SEQ ID NO:101), Figure 61 (SEQ ID NO:103), Figure 63 (SEQ ID NO:110), Figure 65 (SEQ ID NO:115), Figure 67 (SEQ ID NO:117), Figure 69 (SEQ ID NO:122), Figure 71 (SEQ ID NO:127), Figure 73 (SEQ ID NO:129), Figure 75 (SEQ ID NO:131), Figure 77 (SEQ ID NO:133), Figure 79 (SEQ ID NO:135), Figure 81 (SEQ ID NO:137), Figure 83 (SEQ ID NO:139), Figure 85 (SEQ ID NO:141), Figure 87 (SEQ ID NO:143), Figure 89 (SEQ ID NO:145), Figure 91 (SEQ ID NO:147), Figure 93 (SEQ ID NO:152), Figure 95 (SEQ ID NO:157), Figure 97 (SEQ ID NO:159), Figure 99 (SEQ ID NO:161), Figure 101 (SEQ ID NO:169), Figure 103 (SEQ ID NO:179), Figure 105 (SEQ ID NO:188), Figure 107 (SEQ ID NO:193), Figure 109 (SEQ ID NO:195), Figure 111 (SEQ ID NO:197), Figure 113 (SEQ ID NO:202), Figure 115 (SEQ ID NO:209), Figure 117 (SEQ ID NO:211), Figure 119 (SEQ ID NO:213), Figure 121 (SEQ ID NO:215), Figure 123 (SEQ ID NO:217), Figure 125 (SEQ ID NO:219), Figure 127 (SEQ ID NO:224), Figure 129 (SEQ ID NO:226), Figure 131 (SEQ ID NO:228), Figure 133 (SEQ ID NO:233), Figure 135 (SEQ ID NO:235), Figure 137 (SEQ ID NO:242), Figure 139 (SEQ ID NO:247), Figure 141 (SEQ ID NO:252), Figure 143 (SEQ ID NO:259), Figure 145 (SEQ ID NO:264), Figure 147 (SEQ ID NO:266), Figure 149 (SEQ ID NO:268), Figure 151 (SEQ ID NO:270), Figure 153 (SEQ ID NO:272), Figure 155 (SEQ ID NO:274), Figure 157 (SEQ ID NO:276), Figure 159 (SEQ ID NO:281), Figure 161 (SEQ ID NO:286), Figure 163 (SEQ ID NO:291), Figure 165 (SEQ ID NO:296), Figure 167 (SEQ ID NO:301), Figure 169 (SEQ ID NO:303), Figure 171 (SEQ ID NO:305), Figure 173 (SEQ ID NO:307), Figure 175 (SEQ ID NO:309), Figure 177 (SEQ ID NO:314), Figure 179 (SEQ ID NO:316), Figure 181 (SEQ ID NO:321), Figure 183 (SEQ ID NO:323), Figure 185 (SEQ ID NO:325), Figure 187 (SEQ ID NO:327), Figure 189 (SEQ ID NO:329), Figure 191 (SEQ ID NO:331), Figure 193 (SEQ ID NO:333), Figure 195 (SEQ ID NO:335), Figure 197 (SEQ ID NO:337), Figure 199 (SEQ ID NO:339), Figure 201 (SEQ ID NO:346), Figure 203 (SEQ ID NO:351), Figure 205 (SEQ ID NO:353), Figure 207 (SEQ ID NO:355), Figure 209 (SEQ ID NO:357), Figure 211 (SEQ ID NO:363), Figure 213 (SEQ ID NO:365), Figure 215 (SEQ ID NO:371), Figure 217 (SEQ ID NO:373), Figure 219 (SEQ ID NO:375), Figure 221 (SEQ ID NO:377), Figure 223 (SEQ ID NO:382), Figure 225 (SEQ ID NO:384), Figure 227 (SEQ ID NO:389), Figure 229 (SEQ ID NO:394), Figure 231 (SEQ ID NO:396), Figure 233 (SEQ ID NO:401), Figure 235 (SEQ ID NO:405), Figure 237 (SEQ ID NO:409), Figure 239 (SEQ ID NO:414), Figure 241 (SEQ ID NO:422), Figure 242 (SEQ ID NO:428) and Figure 245 (SEQ ID NO:430).

494

4.   Isolated nucleic acid which comprises the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 2.

5.   A vector comprising the nucleic acid of Claim 1.

6.   The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

7.   A host cell comprising the vector of Claim 5.

8.   The host cell of Claim 7 wherein said cell is a CHO cell.

9.   The host cell of Claim 7 wherein said cell is an *E. coli*.

10.   The host cell of Claim 7 wherein said cell is a yeast cell.

11.   A process for producing a PRO polypeptides comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.

12.   Isolated PRO polypeptide having at least 80% sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:4), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 8 (SEQ ID NO:10), Figure 10 (SEQ ID NO:12), Figure 12 (SEQ ID NO:17), Figure 14 (SEQ ID NO:22), Figure 16 (SEQ ID NO:24), Figure 18 (SEQ ID NO:29), Figure 20 (SEQ ID NO:31), Figure 22 (SEQ ID NO:33), Figure 24 (SEQ ID NO:41), Figure 26 (SEQ ID NO:43), Figure 28 (SEQ ID NO:50), Figure 30 (SEQ ID NO:52), Figure 32 (SEQ ID NO:54), Figure 34 (SEQ ID NO:56), Figure 36 (SEQ ID NO:58), Figure 38 (SEQ ID NO:63), Figure 40 (SEQ ID NO:68), Figure 42 (SEQ ID NO:70), Figure 44 (SEQ ID NO:72), Figure 46 (SEQ ID NO:77), Figure 48 (SEQ ID NO:79), Figure 50 (SEQ ID NO:84), Figure 52 (SEQ ID NO:86), Figure 54 (SEQ ID NO:88), Figure 56 (SEQ ID NO:95), Figure 58 (SEQ ID NO:100), Figure 60 (SEQ ID NO:102), Figure 62 (SEQ ID NO:104), Figure 64 (SEQ ID NO:111), Figure 66 (SEQ ID NO:116), Figure 68 (SEQ ID NO:118), Figure 70 (SEQ ID NO:123), Figure 72 (SEQ ID NO:128), Figure 74 (SEQ ID NO:130), Figure 76 (SEQ ID NO:132), Figure 78 (SEQ ID NO:134), Figure 80 (SEQ ID NO:136), Figure 82 (SEQ ID NO:138), Figure 84 (SEQ ID NO:140), Figure 86 (SEQ ID NO:142), Figure 88 (SEQ ID NO:144), Figure 90 (SEQ ID NO:146), Figure 92 (SEQ ID NO:148), Figure 94 (SEQ ID NO:153), Figure 96 (SEQ ID NO:158), Figure 98 (SEQ ID NO:160), Figure 100 (SEQ ID NO:162), Figure 102 (SEQ ID NO:170), Figure 104 (SEQ ID NO:180), Figure 106 (SEQ ID NO:189), Figure 108 (SEQ ID NO:194), Figure 110 (SEQ ID NO:196), Figure 112 (SEQ ID NO:198), Figure 114 (SEQ ID NO:203), Figure 116 (SEQ ID NO:210), Figure 118 (SEQ ID NO:212), Figure 120 (SEQ ID NO:214), Figure 122 (SEQ ID

495

NO:216), Figure 124 (SEQ ID NO:218), Figure 126 (SEQ ID NO:220), Figure 128 (SEQ ID NO:225), Figure 130 (SEQ ID NO:227), Figure 132 (SEQ ID NO:229), Figure 134 (SEQ ID NO:234), Figure 136 (SEQ ID NO:236), Figure 138 (SEQ ID NO:243), Figure 140 (SEQ ID NO:248), Figure 142 (SEQ ID NO:253), Figure 144 (SEQ ID NO:260), Figure 146 (SEQ ID NO:265), Figure 148 (SEQ ID NO:267), Figure 150 (SEQ ID NO:269), Figure 152 (SEQ ID NO:271), Figure 154 (SEQ ID NO:273), Figure 156 (SEQ ID NO:275), Figure

5   158 (SEQ ID NO:277), Figure 160 (SEQ ID NO:282), Figure 162 (SEQ ID NO:287), Figure 164 (SEQ ID NO:292), Figure 166 (SEQ ID NO:297), Figure 168 (SEQ ID NO:302), Figure 170 (SEQ ID NO:304), Figure 172 (SEQ ID NO:306), Figure 174 (SEQ ID NO:308), Figure 176 (SEQ ID NO:310), Figure 178 (SEQ ID NO:315), Figure 180 (SEQ ID NO:317), Figure 182 (SEQ ID NO:322), Figure 184 (SEQ ID NO:324), Figure 186 (SEQ ID NO:326), Figure 188 (SEQ ID NO:328), Figure 190 (SEQ ID NO:330), Figure 192 (SEQ ID

10   NO:332), Figure 194 (SEQ ID NO:334), Figure 196 (SEQ ID NO:336), Figure 198 (SEQ ID NO:338), Figure 200 (SEQ ID NO:340), Figure 202 (SEQ ID NO:347), Figure 204 (SEQ ID NO:352), Figure 206 (SEQ ID NO:354), Figure 208 (SEQ ID NO:356), Figure 210 (SEQ ID NO:358), Figure 212 (SEQ ID NO:364), Figure 214 (SEQ ID NO:366), Figure 216 (SEQ ID NO:372), Figure 218 (SEQ ID NO:374), Figure 220 (SEQ ID NO:376), Figure 222 (SEQ ID NO:378), Figure 224 (SEQ ID NO:383), Figure 226 (SEQ ID NO:385), Figure

15   228 (SEQ ID NO:390), Figure 230 (SEQ ID NO:395), Figure 232 (SEQ ID NO:397), Figure 234 (SEQ ID NO:402), Figure 236 (SEQ ID NO:406), Figure 238 (SEQ ID NO:410), Figure 240 (SEQ ID NO:415), Figure 242 (SEQ ID NO:423), Figure 244 (SEQ ID NO:429) and Figure 246 (SEQ ID NO:431).

20   13.   Isolated PRO polypeptide having at least 80% sequence identity to the amino acid sequence encoded by a nucleic acid molecule deposited under any ATCC accession number shown in Table 2.

14.   A chimeric molecule comprising a polypeptide according to Claim 12 fused to a heterologous amino acid sequence.

25   15.   The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is an epitope tag sequence.

16.   The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

30

17.   An antibody which specifically binds to a PRO polypeptide according to Claim 12.

18.   The antibody of Claim 17 wherein said antibody is a monoclonal antibody.

35   19.   The antibody of Claim 17 wherein said antibody is a humanized antibody.

20.   The antibody of Claim 17 wherein said antibody is an antibody fragment.

496

21.   An isolated nucleic acid which has at least 80% sequence identity to a nucleic acid sequence
selected from the group consisting of that shown in Figure 1 (SEQ ID NO:3), Figure 3 (SEQ ID NO:5), Figure
5 (SEQ ID NO:7), Figure 7 (SEQ ID NO:9), Figure 9 (SEQ ID NO:11), Figure 11 (SEQ ID NO:16), Figure
13 (SEQ ID NO:21), Figure 15 (SEQ ID NO:23), Figure 17 (SEQ ID NO:28), Figure 19 (SEQ ID NO:30),
Figure 21 (SEQ ID NO:32), Figure 23 (SEQ ID NO:40), Figure 25 (SEQ ID NO:42), Figure 27 (SEQ ID
NO:49), Figure 29 (SEQ ID NO:51), Figure 31 (SEQ ID NO:53), Figure 33 (SEQ ID NO:55), Figure 35 (SEQ
ID NO:57), Figure 37 (SEQ ID NO:62), Figure 39 (SEQ ID NO:67), Figure 41 (SEQ ID NO:69), Figure 43
(SEQ ID NO:71), Figure 45 (SEQ ID NO:76), Figure 47 (SEQ ID NO:78), Figure 49 (SEQ ID NO:83), Figure
51 (SEQ ID NO:85), Figure 53 (SEQ ID NO:87), Figure 55 (SEQ ID NO:94), Figure 57 (SEQ ID NO:99),
Figure 59 (SEQ ID NO:101), Figure 61 (SEQ ID NO:103), Figure 63 (SEQ ID NO:110), Figure 65 (SEQ ID
NO:115), Figure 67 (SEQ ID NO:117), Figure 69 (SEQ ID NO:122), Figure 71 (SEQ ID NO:127), Figure 73
(SEQ ID NO:129), Figure 75 (SEQ ID NO:131), Figure 77 (SEQ ID NO:133), Figure 79 (SEQ ID NO:135),
Figure 81 (SEQ ID NO:137), Figure 83 (SEQ ID NO:139), Figure 85 (SEQ ID NO:141), Figure 87 (SEQ ID
NO:143), Figure 89 (SEQ ID NO:145), Figure 91 (SEQ ID NO:147), Figure 93 (SEQ ID NO:152), Figure 95
(SEQ ID NO:157), Figure 97 (SEQ ID NO:159), Figure 99 (SEQ ID NO:161), Figure 101 (SEQ ID NO:169),
Figure 103 (SEQ ID NO:179), Figure 105 (SEQ ID NO:188), Figure 107 (SEQ ID NO:193), Figure 109 (SEQ
ID NO:195), Figure 111 (SEQ ID NO:197), Figure 113 (SEQ ID NO:202), Figure 115 (SEQ ID NO:209),
Figure 117 (SEQ ID NO:211), Figure 119 (SEQ ID NO:213), Figure 121 (SEQ ID NO:215), Figure 123 (SEQ
ID NO:217), Figure 125 (SEQ ID NO:219), Figure 127 (SEQ ID NO:224), Figure 129 (SEQ ID NO:226),
Figure 131 (SEQ ID NO:228), Figure 133 (SEQ ID NO:233), Figure 135 (SEQ ID NO:235), Figure 137 (SEQ
ID NO:242), Figure 139 (SEQ ID NO:247), Figure 141 (SEQ ID NO:252), Figure 143 (SEQ ID NO:259),
Figure 145 (SEQ ID NO:264), Figure 147 (SEQ ID NO:266), Figure 149 (SEQ ID NO:268), Figure 151 (SEQ
ID NO:270), Figure 153 (SEQ ID NO:272), Figure 155 (SEQ ID NO:274), Figure 157 (SEQ ID NO:276),
Figure 159 (SEQ ID NO:281), Figure 161 (SEQ ID NO:286), Figure 163 (SEQ ID NO:291), Figure 165 (SEQ
ID NO:296), Figure 167 (SEQ ID NO:301), Figure 169 (SEQ ID NO:303), Figure 171 (SEQ ID NO:305),
Figure 173 (SEQ ID NO:307), Figure 175 (SEQ ID NO:309), Figure 177 (SEQ ID NO:314), Figure 179 (SEQ
ID NO:316), Figure 181 (SEQ ID NO:321), Figure 183 (SEQ ID NO:323), Figure 185 (SEQ ID NO:325),
Figure 187 (SEQ ID NO:327), Figure 189 (SEQ ID NO:329), Figure 191 (SEQ ID NO:331), Figure 193 (SEQ
ID NO:333), Figure 195 (SEQ ID NO:335), Figure 197 (SEQ ID NO:337), Figure 199 (SEQ ID NO:339),
Figure 201 (SEQ ID NO:346), Figure 203 (SEQ ID NO:351), Figure 205 (SEQ ID NO:353), Figure 207 (SEQ
ID NO:355), Figure 209 (SEQ ID NO:357), Figure 211 (SEQ ID NO:363), Figure 213 (SEQ ID NO:365),
Figure 215 (SEQ ID NO:371), Figure 217 (SEQ ID NO:373), Figure 219 (SEQ ID NO:375), Figure 221 (SEQ
ID NO:377), Figure 223 (SEQ ID NO:382), Figure 225 (SEQ ID NO:384), Figure 227 (SEQ ID NO:389),
Figure 229 (SEQ ID NO:394), Figure 231 (SEQ ID NO:396), Figure 233 (SEQ ID NO:401), Figure 235 (SEQ
ID NO:405), Figure 237 (SEQ ID NO:409), Figure 239 (SEQ ID NO:414), Figure 241 (SEQ ID NO:422),
Figure 242 (SEQ ID NO:428) and Figure 245 (SEQ ID NO:430).

22.   An isolated nucleic acid which has at least 80% sequence identity to the full-length coding sequence of a nucleotide sequence selected from the group consisting of that shown in Figure 1 (SEQ ID NO:3), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 7 (SEQ ID NO:9), Figure 9 (SEQ ID NO:11), Figure 11 (SEQ ID NO:16), Figure 13 (SEQ ID NO:21), Figure 15 (SEQ ID NO:23), Figure 17 (SEQ ID NO:28), Figure 19 (SEQ ID NO:30), Figure 21 (SEQ ID NO:32), Figure 23 (SEQ ID NO:40), Figure 25 (SEQ ID NO:42), Figure 27 (SEQ ID NO:49), Figure 29 (SEQ ID NO:51), Figure 31 (SEQ ID NO:53), Figure 33 (SEQ ID NO:55), Figure 35 (SEQ ID NO:57), Figure 37 (SEQ ID NO:62), Figure 39 (SEQ ID NO:67), Figure 41 (SEQ ID NO:69), Figure 43 (SEQ ID NO:71), Figure 45 (SEQ ID NO:76), Figure 47 (SEQ ID NO:78), Figure 49 (SEQ ID NO:83), Figure 51 (SEQ ID NO:85), Figure 53 (SEQ ID NO:87), Figure 55 (SEQ ID NO:94), Figure 57 (SEQ ID NO:99), Figure 59 (SEQ ID NO:101), Figure 61 (SEQ ID NO:103), Figure 63 (SEQ ID NO:110), Figure 65 (SEQ ID NO:115), Figure 67 (SEQ ID NO:117), Figure 69 (SEQ ID NO:122), Figure 71 (SEQ ID NO:127), Figure 73 (SEQ ID NO:129), Figure 75 (SEQ ID NO:131), Figure 77 (SEQ ID NO:133), Figure 79 (SEQ ID NO:135), Figure 81 (SEQ ID NO:137), Figure 83 (SEQ ID NO:139), Figure 85 (SEQ ID NO:141), Figure 87 (SEQ ID NO:143), Figure 89 (SEQ ID NO:145), Figure 91 (SEQ ID NO:147), Figure 93 (SEQ ID NO:152), Figure 95 (SEQ ID NO:157), Figure 97 (SEQ ID NO:159), Figure 99 (SEQ ID NO:161), Figure 101 (SEQ ID NO:169), Figure 103 (SEQ ID NO:179), Figure 105 (SEQ ID NO:188), Figure 107 (SEQ ID NO:193), Figure 109 (SEQ ID NO:195), Figure 111 (SEQ ID NO:197), Figure 113 (SEQ ID NO:202), Figure 115 (SEQ ID NO:209), Figure 117 (SEQ ID NO:211), Figure 119 (SEQ ID NO:213), Figure 121 (SEQ ID NO:215), Figure 123 (SEQ ID NO:217), Figure 125 (SEQ ID NO:219), Figure 127 (SEQ ID NO:224), Figure 129 (SEQ ID NO:226), Figure 131 (SEQ ID NO:228), Figure 133 (SEQ ID NO:233), Figure 135 (SEQ ID NO:235), Figure 137 (SEQ ID NO:242), Figure 139 (SEQ ID NO:247), Figure 141 (SEQ ID NO:252), Figure 143 (SEQ ID NO:259), Figure 145 (SEQ ID NO:264), Figure 147 (SEQ ID NO:266), Figure 149 (SEQ ID NO:268), Figure 151 (SEQ ID NO:270), Figure 153 (SEQ ID NO:272), Figure 155 (SEQ ID NO:274), Figure 157 (SEQ ID NO:276), Figure 159 (SEQ ID NO:281), Figure 161 (SEQ ID NO:286), Figure 163 (SEQ ID NO:291), Figure 165 (SEQ ID NO:296), Figure 167 (SEQ ID NO:301), Figure 169 (SEQ ID NO:303), Figure 171 (SEQ ID NO:305), Figure 173 (SEQ ID NO:307), Figure 175 (SEQ ID NO:309), Figure 177 (SEQ ID NO:314), Figure 179 (SEQ ID NO:316), Figure 181 (SEQ ID NO:321), Figure 183 (SEQ ID NO:323), Figure 185 (SEQ ID NO:325), Figure 187 (SEQ ID NO:327), Figure 189 (SEQ ID NO:329), Figure 191 (SEQ ID NO:331), Figure 193 (SEQ ID NO:333), Figure 195 (SEQ ID NO:335), Figure 197 (SEQ ID NO:337), Figure 199 (SEQ ID NO:339), Figure 201 (SEQ ID NO:346), Figure 203 (SEQ ID NO:351), Figure 205 (SEQ ID NO:353), Figure 207 (SEQ ID NO:355), Figure 209 (SEQ ID NO:357), Figure 211 (SEQ ID NO:363), Figure 213 (SEQ ID NO:365), Figure 215 (SEQ ID NO:371), Figure 217 (SEQ ID NO:373), Figure 219 (SEQ ID NO:375), Figure 221 (SEQ ID NO:377), Figure 223 (SEQ ID NO:382), Figure 225 (SEQ ID NO:384), Figure 227 (SEQ ID NO:389), Figure 229 (SEQ ID NO:394), Figure 231 (SEQ ID NO:396), Figure 233 (SEQ ID NO:401), Figure 235 (SEQ ID NO:405), Figure 237 (SEQ ID NO:409), Figure 239 (SEQ ID NO:414), Figure 241 (SEQ ID NO:422), Figure 242 (SEQ ID NO:428) and Figure 245 (SEQ ID NO:430).

23.   An isolated extracellular domain of a PRO polypeptide.

498

24.   An isolated PRO polypeptide lacking its associated signal peptide.

25.   An isolated polypeptide having at least about 80% amino acid sequence identity to an extracellular domain of of PRO polypeptide.

5   26.   An isolated polypeptide having at least about 80% amino acid sequence identity to a PRO polypeptide lacking its associated signal peptide.

27.   An isolated nucleic acid encoding the polypeptide of any one of Claims 23 to 26.

FIGURE 1

CCAATCGCCCGGTGCGGTGGTGCAGGGTCTCGGGCTAGTCATGGCGTCCCCGTCTCGGAGAC

TGCAGACTAAACCAGTCATTACTTGTTTCAAGAGCGTTCTGCTAATCTACACTTTTATTTTC

TGGATCACTGGCGTTATCCTTCTTGCAGTTGGCATTTGGGGCAAGGTGAGCCTGGAGAATTA

CTTTTCTCTTTTAAATGAGAAGGCCACCAATGTCCCCTTCGTGCTCATTGCTACTGGTACCG

TCATTATTCTTTTGGGCACCTTTGGTTGTTTTGCTACCTGCCGAGCTTCTGCATGGATGCTA

AAACTGTATGCAATGTTTCTGACTCTCGTTTTTTTTGGTCGAACTGGTCGCTGCCATCGTAGG

ATTTGTTTTCAGACATGAGATTAAGAACAGCTTTAAGAATAATTATGAGAAGGCTTTGAAGC

AGTATAACTCTACAGGAGATTATAGAAGCCATGCAGTAGACAAGATCCAAAATACGTTGCAT

TGTTGTGGTGTCACCGATTATAGAGATTGGACAGATACTAATTATTACTCAGAAAAAGGATT

TCCTAAGAGTTGCTGTAAACTTGAAGATTGTACTCCACAGAGAGATGCAGACAAAGTAAACA

ATGAAGGTTGTTTTATAAAGGTGATGACCATTATAGAGTCAGAAATGGGAGTCGTTGCAGGA

ATTTCCTTTGGAGTTGCTTGCTTCCAACTGATTGGAATCTTTCTCGCCTACTGCCWCTCTCG

TGCCATAACAAATAACCAGTATGAGATAGTGTAACCCAATGTATCTGTGGGCCTATTCCTCT

CTACCTTTAAGGACATTTAGGGTCCCCCCTGTGAATTAGAAAGTTGCTTGGCTGGAGAACTG

ACAACACTACTTACTGATAGACCAAAAAACTACACCAGTAGGTTGATTCAATCAAGATGTAT

GTAGACCTAAAACTACACCAATAGGCTGATTCAATCAAGATCCGTGCTCGCAGTGGGCTGAT

TCAATCAAGATGTATGTTTGCTATGTTCTAAGTCCACCTTCTATCCCATTCATGTTAGATCG

TTGAAACCCTGTATCCCTCTGAAACACTGGAAGAGCTAGTAAATTGTAAATGAAGT

FIGURE 2

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA19902

><subunit 1 of 1, 245 aa, 1 stop, 1 unknown

><MW: -1, pI: 8.36, NX(S/T): 1

MASPSRRLQTKPVITCFKSVLLIYTFIFWITGVILLAVGIWGKVSLENYFSLLNEKATNVPF

VLIATGTVIILLGTFGCFATCRASAWMLKLYAMFLTLVFLVELVAAIVGFVFRHEIKNSFKN

NYEKALKQYNSTGDYRSHAVDKIQNTLHCCGVTDYRDWTDTNYYSEKGFPKSCCKLEDCTPQ

RDADKVNNEGCFIKVMTIIESEMGVVAGISFGVACFQLIGIFLAYCXSRAITNNQYEIV

Important features of the protein:

Signal peptide:

amino acids 1-42

Transmembrane domains:

amino acids 19-42, 61-83, 92-114, 209-230,

N-glycosylation site.

amino acids 134-138

Tyrosine kinase phosphorylation site.

amino acids 160-168, 160-169

N-myristoylation site.

amino acids 75-81, 78-84, 210-216, 214-220, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 69-80, 211-222

FIGURE 3

CCCACGCGTCCGGCGCCGTGGCCTCGCGTCCATCTTTGCCGTTCTCTCGGACCTGTCACAAA

GGAGTCGCGCCGCCGCCGCCGCCCCCTCCCTCCGGTGGGCCCGGGAGGTAGAGAAAGTCAGT

GCCACAGCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCTGAGGGT

TGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACCGGACTGGG

GCCGGGGTAGGCTCTGGAAAGGGCCCGGGAGAGAGGTGGCGTTGGTCAGAACCTGAGAAACA

GCCGAGAGGTTTTCCACCGAGGCCCGCGCTTGAGGGATCTGAAGAGGTTCCTAGAAGAGGGT

GTTCCCTCTTTCGGGGGTCCTCACCAGAAGAGGTTCTTGGGGGTCGCCCTTCTGAGGAGGCT

GCGGCTAACAGGGCCCAGAACTGCCATTGGATGTCCAGAATCCCCTGTAGTTGATAATGTTG

GGAATAAGCTCTGCAACTTTCTTTGGCATTCAGTTGTTAAAAACAAATAGGATGCAAATTCC

TCAACTCCAGGTTATGAAAACAGTACTTGGAAAACTGAAAACTACCTAA**ATG**ATCGTCTTTG

GTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGGCCAGGGTCTGTTGTTGACTCTCGAAGAG

CACATAGCCCACTTCCTAGGGACTGGAGGTGCCGCTACTACCATGGGTAATTCCTGTATCTG

CCGAGATGACAGTGGAACAGATGACAGTGTTGACACCCAACAGCAACAGGCCGAGAACAGTG

CAGTACCCACTGCTGACACAAGGAGCCAACCACGGGACCCTGTTCGGCCACCAAGGAGGGGC

CGAGGACCTCATGAGCCAAGGAGAAAGAAACAAAATGTGGATGGGCTAGTGTTGGACACACT

GGCAGTAATACGGACTCTTGTAGATAAG**TAA**GTATCTGACTCACGGTCACCTCCAGTGGAAT

GAAAAGTGTTCTGCCCGGAACCATGACTTTAGGACTCCTTCAGTTCCTTTAGGACATACTCG

CCAAGCCTTGTGCTCACAGGGCAAAGGAGAATATTTTAATGCTCCGCTGATGGCAGAGTAAA

TGATAAGATTTGATGTTTTTGCTTGCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTC

TGTAGCAGAAAACACGATAAAGCTATGATCTTTATTAGAG

FIGURE 4

MIVFGWAVFLASRSLGQGLLLTLEEHIAHFLGTGGAATTMGNSCICRDDSGTDDSVDTQQQQ

AENSAVPTADTRSQPRDPVRPPRRGRGPHEPRRKKQNVDGLVLDTLAVIRTLVDKO

Signal peptide:

amino acids 1-16

Casein kinase II phosphorylation site.

amino acids 22-26, 50-54, 113-117

N-myristoylation site.

amino acids 18-24, 32-38, 34-40, 35-41, 51-57

FIGURE 5

GGCACGAGGCGCTGTCCACCCGGGGGCGTGGGAGTGAGGTACCAGATTCAGCCCATTTGGCC

CCGACGCCTCTGTTCTCGGAATCCGGGTGCTGCGGATTGAGGTCCCGGTTCCTAACGGACTG

CAAG**ATG**GAGGAAGGCGGGAACCTAGGAGGCCTGATTAAGATGGTCCATCTACTGGTCTTGT

CAGGTGCCTGGGGCATGCAAATGTGGGTGACCTTCGTCTCAGGCTTCCTGCTTTTCCGAAGC

CTTCCCCGACATACCTTCGGACTAGTGCAGAGCAAACTCTTCCCCTTCTACTTCCACATCTC

CATGGGCTGTGCCTTCATCAACCTCTGCATCTTGGCTTCACAGCATGCTTGGGCTCAGCTCA

CATTCTGGGAGGCCAGCCAGCTTTACCTGCTGTTCCTGAGCCTTACGCTGGCCACTGTCAAC

GCCCGCTGGCTGGAACCCCGCACCACAGCTGCCATGTGGGCCCTGCAAACCGTGGAGAAGGA

GCGAGGCCTGGGTGGGGAGGTACCAGGCAGCCACCAGGGTCCCGATCCCTACCGCCAGCTGC

GAGAGAAGGACCCCAAGTACAGTGCTCTCCGCCAGAATTTCTTCCGCTACCATGGGCTGTCC

TCTCTTTGCAATCTGGGCTGCGTCCTGAGCAATGGGCTCTGTCTCGCTGGCCTTGCCCTGGA

AATAAGGAGCCTC**TAG**CATGGGCCCTGCATGCTAATAAATGCTTCTTCAGAAATGAAAAAAA

AAAAAAAAAAAA

FIGURE 6

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56107
<subunit 1 of 1, 231 aa, 1 stop
<NX(S/T): 0
MEEGGNLGGLIKMVHLLVLSGAWGMQMWVTFVSGFLLFRSLPRHTFGLVQSKLFPFYFHISM
GCAFINLCILASQHAWAQLTFWEASQLYLLFLSLTLATVNARWLEPRTTAAMWALQTVEKER
GLGGEVPGSHQGPDPYRQLREKDPKYSALRQNFFRYHGLSSLCNLGCVLSNGLCLAGLALEIRSL
```

Signal peptide:
amino acids 1-24

Transmembrane domain:
amino acids 86-103, 60-75

Casein kinase II phosphorylation site.
amino acids 82-86

Tyrosine kinase phosphorylation site.
amino acids 144-151

N-myristoylation site.
amino acids 4-10, 5-11, 47-53, 170-176, 176-182

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 54-65

G-protein coupled receptors proteins.
amino acids 44-85

6/270

FIGURE 7

AATTCAGATTTTAAGCCCATTCTGCAGTGGAATTTCATGAACTAGCAAGAGGACACCATCTT

CTTGTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAA**ATG**CTCTTTTGGGT

GCTAGGCCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAG

ACATCACTGATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAAACTTGGCAGCC

AGAACTTTTGATAAAAAGGGATTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAAC

AGCTTTAAAGGCAGAAACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAG

AGAATGTCAAGAGGACTGCCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGT

CTGATCAATAATGCTGGTGTTCCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGA

CTACAGAGAACCTATTGAAGTGAACCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTC

CTTTGGTCAAGAAAGCTCAAGGGAGAGTTATTAATGTCTCCAGTGTTGGAGGTCGCCTTGCA

ATCGTTGGAGGGGGCTATACTCCATCCAAATATGCAGTGGAAGGTTTCAATGACAGCTTAAG

ACGGGACATGAAAGCTTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAAACAA

ACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAACTCGCCATTTGGGAGCAGCTGTCTCCA

GACATCAAACAACAATATGGAGAAGGTTACATTGAAAAAAAGTCTAGACAAACTGAAAGGCAA

TAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTGCATGGACCACGCTCTAACAA

GTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAATTTTCTGGATACCTCTG

TCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGGCTAA

TCCCAAGGCAGTG**TGA**CTCAGCTAACCACAAATGTCTCCTCCAGGCTATGAAATTGGCCGAT

TTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACTCATTTAGA

TCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGTCCCTG

CTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCTGT

ATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATG

ATCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTT

AAGTATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAA

FIGURE 8

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56406
><subunit 1 of 1, 319 aa, 1 stop
><MW: 35227, pI: 8.97, NX(S/T): 3
MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLT
ESGSTALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDW
LTLEDYREPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGF
NDSLRRDMKAFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLD
KLKGNKSYVNMDLSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQK
AELANPKAV
```

Important features of the protein:
Signal peptide:
amino acids 1-17

Transmembrane domain:
amino acids 136-152

N-glycosylation sites.
amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.
amino acids 39-42

N-myristoylation sites.
amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 9

```
GCGGGCTGTTGACGGCGCTGCGATGGCTGCCTGCGAGGGCAGGAGAAGCGGAGCTCTCGGTT
CCTCTCAGTCGGACTTCCTGACGCCGCCAGTGGGCGGGGCCCCTTGGGCCGTCGCCACCACT
GTAGTCATGTACCCACCGCCGCCGCCGCCGCCTCATCGGGACTTCATCTCGGTGACGCTGAG
CTTTGGCGAGAGCTATGACAACAGCAAGAGTTGGCGGCGGCGCTCGTGCTGGAGGAAATGGA
AGCAACTGTCGAGATTGCAGCGGAATATGATTCTCTTCCTCCTTGCCTTTCTGCTTTTCTGT
GGACTCCTCTTCTACATCAACTTGGCTGACCATTGGAAAGCTCTGGCTTTCAGGCTAGAGGA
AGAGCAGAAGATGAGGCCAGAAATTGCTGGGTTAAAACCAGCAAATCCACCCGTCTTACCAG
CTCCTCAGAAGGCGGACACCGACCCTGAGAACTTACCTGAGATTTCGTCACAGAAGACACAA
AGACACATCCAGCGGGGACCACCTCACCTGCAGATTAGACCCCCAAGCCAAGACCTGAAGGA
TGGGACCCAGGAGGAGGCCACAAAAAGGCAAGAAGCCCCTGTGGATCCCCGCCCGGAAGGAG
ATCCGCAGAGGACAGTCATCAGCTGGAGGGGAGCGGTGATCGAGCCTGAGCAGGGCACCGAG
CTCCCTTCAAGAAGAGCAGAAGTGCCCACCAAGCCTCCCCTGCCACCGGCCAGGACACAGGG
CACACCAGTGCATCTGAACTATCGCCAGAAGGGCGTGATTGACGTCTTCCTGCATGCATGGA
AAGGATACCGCAAGTTTGCATGGGGCCATGACGAGCTGAAGCCTGTGTCCAGGTCCTTCAGT
GAGTGGTTTGGCCTCGGTCTCACACTGATCGACGCGCTGGACACCATGTGGATCTTGGGTCT
GAGGAAAGAATTTGAGGAAGCCAGGAAGTGGGTGTCGAAGAAGTTACACTTTGAAAAGGACG
TGGACGTCAACCTGTTTGAGAGCACGATCCGCATCCTGGGGGGGCTCCTGAGTGCCTACCAC
CTGTCTGGGGACAGCCTCTTCCTGAGGAAAGCTGAGGATTTTGGAAATCGGCTAATGCCTGC
CTTCAGAACACCATCCAAGATTCCTTACTCGGATGTGAACATCGGTACTGGAGTTGCCCACC
CGCCACGGTGGACCTCCGACAGCACTGTGGCCGAGGTGACCAGCATTCAGCTGGAGTTCCGG
GAGCTCTCCCGTCTCACAGGGGATAAGAAGTTTCAGGAGGCAGTGGAGAAGGTGACACAGCA
CATCCACGGCCTGTCTGGGAAGAAGGATGGGCTGGTGCCCATGTTCATCAATACCCACAGTG
GCCTCTTCACCCACCTGGGCGTATTCACGCTGGGCGCCAGGGCCGACAGCTACTATGAGTAC
CTGCTGAAGCAGTGGATCCAGGGCGGGAAGCAGGAGACACAGCTGCTGGAAGACTACGTGGA
AGCCATCGAGGGTGTCAGAACGCACCTGCTGCGGCACTCCGAGCCCAGTAAGCTCACCTTTG
TGGGGGAGCTTGCCCACGGCCGCTTCAGTGCCAAGATGGACCACCTGGTGTGCTTCCTGCCA
GGGACGCTGGCTCTGGGCGTCTACCACGGCCTGCCCGCCAGCCACATGGAGCTGGCCCAGGA
GCTCATGGAGACTTGTTACCAGATGAACCGGCAGATGGAGACGGGGCTGAGTCCCGAGATCG
TGCACTTCAACCTTTACCCCCAGCCGGGCCGTCGGGACGTGGAGGTCAAGCCAGCAGACAGG
CACAACCTGCTGCGGCCAGAGACCGTGGAGAGCCTGTTCTACCTGTACCGCGTCACAGGGGA
CCGCAAATACCAGGACTGGGGCTGGGAGATTCTGCAGAGCTTCAGCCGATTCACACGGGTCC
CCTCGGGTGGCTATTCTTCCATCAACAATGTCCAGGATCCTCAGAAGCCCGAGCCTAGGGAC
AAGATGGAGAGCTTCTTCCTGGGGGAGACGCTCAAGTATCTGTTCTTGCTCTTCTCCGATGA
CCCAAACCTGCTCAGCCTGGACGCCTACGTGTTCAACACCGAAGCCCACCCTCTGCCTATCT
GGACCCCTGCCTAGGGTGGATGGCTGCTGGTGTGGGGACTTCGGGTGGGCAGAGGCACCTTG
CTGGGTCTGTGGCATTTTCCAAGGGCCCACGTAGCACCGGCAACCGCCAAGTGGCCCAGGCT
CTGAACTGGCTCTGGGCTCCTCCTCGTCTCTGCTTTAATCAGGACACCGTGAGGACAAGTGA
GGCCGTCAGTCTTGGTGTGATGCGGGGTGGGCTGGGCCGCTGGAGCCTCCGCCTGCTTCCTC
CAGAAGACACGAATCATGACTCACGATTGCTGAAGCCTGAGCAGGTCTCTGTGGGCCGACCA
GAGGGGGGCTTCGAGGTGGTCCCTGGTACTGGGGTGACCGAGTGGACAGCCCAGGGTGCAGC
TCTGCCCGGGCTCGTGAAGCCTCAGATGTCCCCAATCCAAGGGTCGGAGGGGCTGCCGTGA
CTCCAGAGGCCTGAGGCTCCAGGGCTGGCTCTGGTGTTTACAAGCTGGACTCAGGGATCCTC
CTGGCCGCCCCGCAGGGGGCTTGGAGGGCTGGACGGCAAGTCCGTCTAGCTCACGGGCCCCT
CCAGTGGAATGGGTCTTTTCGGTGGAGATAAAAGTTGATTTGCTCTAACCGCAA
```

FIGURE 10

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56529

><subunit 1 of 1, 699 aa, 1 stop

><MW: 79553, pI: 7.83, NX(S/T): 0

MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVVMYPPPPPPPHRDFISVTLSFGESYDN

SKSWRRRSCWRKWKQLSRLQRNMILFLLAFLLFCGLLFYINLADHWKALAFRLEEEQKMRPE

IAGLKPANPPVLPAPQKADTDPENLPEISSQKTQRHIQRGPPHLQIRPPSQDLKDGTQEEAT

KRQEAPVDPRPEGDPQRTVISWRGAVIEPEQGTELPSRRAEVPTKPPLPPARTQGTPVHLNY

RQKGVIDVFLHAWKGYRKFAWGHDELKPVSRSFSEWFGLGLTLIDALDTMWILGLRKEFEEA

RKWVSKKLHFEKDVDVNLFESTIRILGGLLSAYHLSGDSLFLRKAEDFGNRLMPAFRTPSKI

PYSDVNIGTGVAHPPRWTSDSTVAEVTSIQLEFRELSRLTGDKKFQEAVEKVTQHIHGLSGK

KDGLVPMFINTHSGLFTHLGVFTLGARADSYYEYLLKQWIQGGKQETQLLEDYVEAIEGVRT

HLLRHSEPSKLTFVGELAHGRFSAKMDHLVCFLPGTLALGVYHGLPASHMELAQELMETCYQ

MNRQMETGLSPEIVHFNLYPQPGRRDVEVKPADRHNLLRPETVESLFYLYRVTGDRKYQDWG

WEILQSFSRFTRVPSGGYSSINNVQDPQKPEPRDKMESFFLGETLKYLFLLFSDDPNLLSLD

AYVFNTEAHPLPIWTPA

Important features of the protein:

Transmembrane domain:

amino acids 21-40 and 84-105 (type II)

FIGURE 11

GGCGCCGCGTAGGCCCGGGAGGCCGGGCCGGCCGGGCTGCGAGCGCCTGCCCCATGCGCCGC

CGCCTCTCCGCACG<u>ATG</u>TTCCCCTCGCGGAGGAAAGCGGCGCAGCTGCCCTGGGAGGACGGC

AGGTCCGGGTTGCTCTCCGGCGGCCTCCCTCGGAAGTGTTCCGTCTTCCACCTGTTCGTGGC

CTGCCTCTCGCTGGGCTTCTTCTCCCTACTCTGGCTGCAGCTCAGCTGCTCTGGGGACGTGG

CCCGGGCAGTCAGGGGACAAGGGCAGGAGACCTCGGGCCCTCCCCGTGCCTGCCCCCCAGAG

CCGCCCCCTGAGCACTGGGAAGAAGACGCATCCTGGGGCCCCCACCGCCTGGCAGTGCTGGT

GCCCTTCCGCGAACGCTTCGAGGAGCTCCTGGTCTTCGTGCCCCACATGCGCCGCTTCCTGA

GCAGGAAGAAGATCCGGCACCACATCTACGTGCTCAACCAGGTGGACCACTTCAGGTTCAAC

CGGGCAGCGCTCATCAACGTGGGCTTCCTGGAGAGCAGCAACAGCACGGACTACATTGCCAT

GCACGACGTTGACCTGCTCCCTCTCAACGAGGAGCTGGACTATGGCTTTCCTGAGGCTGGGC

CCTTCCACGTGGCCTCCCCGGAGCTCCACCCTCTCTACCACTACAAGACCTATGTCGGCGGC

ATCCTGCTGCTCTCCAAGCAGCACTACCGGCTGTGCAATGGGATGTCCAACCGCTTCTGGGG

CTGGGGCCGCGAGGACGACGAGTTCTACCGGCGCATTAAGGGAGCTGGGCTCCAGCTTTTCC

GCCCCTCGGGAATCACAACTGGGTACAAGACATTTCGCCACCTGCATGACCCAGCCTGGCGG

AAGAGGGACCAGAAGCGCATCGCAGCTCAAAAACAGGAGCAGTTCAAGGTGGACAGGGAGGG

AGGCCTGAACACTGTGAAGTACCATGTGGCTTCCCGCACTGCCCTGTCTGTGGGCGGGGCCC

CCTGCACTGTCCTCAACATCATGTTGGACTGTGACAAGACCGCCACACCCTGGTGCACATTC

AGC<u>TGA</u>GCTGGATGGACAGTGAGGAAGCCTGTACCTACAGGCCATATTGCTCAGGCTCAGGA

CAAGGCCTCAGGTCGTGGGCCCAGCTCTGACAGGATGTGGAGTGGCCAGGACCAAGACAGCA

AGCTACGCAATTGCAGCCACCCGGCCGCCAAGGCAGGCTTGGGCTGGGCCAGGACACGTGGG

GTGCCTGGGACGCTGCTTGCCATGCACAGTGATCAGAGAGAGGCTGGGGTGTGTCCTGTCCG

GGACCCCCCCTGCCTTCCTGCTCACCCTACTCTGACCTCCTTCACGTGCCCAGGCCTGTGGG

TAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCCTACTCTGACCTCCTTCACGTGCCC

AGGCCTGTGGGTAGTGGGGAGGGCTGAACAGGACAACCTCTCATCACCCCCAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 12

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56531

><subunit 1 of 1, 327 aa, 1 stop

><MW: 37406, pI: 9.30, NX(S/T): 1

MFPSRRKAAQLPWEDGRSGLLSGGLPRKCSVFHLFVACLSLGFFSLLWLQLSCSGDVARAVR

GQGQETSGPPRACPPEPPPEHWEEDASWGPHRLAVLVPFRERFEELLVFVPHMRRFLSRKKI

RHHIYVLNQVDHFRFNRAALINVGFLESSNSTDYIAMHDVDLLPLNEELDYGFPEAGPFHVA

SPELHPLYHYKTYVGGILLLSKQHYRLCNGMSNRFWGWGREDDEFYRRIKGAGLQLFRPSGI

TTGYKTFRHLHDPAWRKRDQKRIAAQKQEQFKVDREGGLNTVKYHVASRTALSVGGAPCTVL

NIMLDCDKTATPWCTFS

Signal peptide:
amino acids 1-42

Transmembrane domain:
amino acids 29-49 (type II)

N-glycosylation site.
amino acids 154-158

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 27-31

Tyrosine kinase phosphorylation site.
amino acids 226-233

N-myristoylation site.
amino acids 19-25, 65-71, 247-253, 285-291, 303-309, 304-310

12/270

FIGURE 13

CAATGTTTGCCTATCCACCTCCCCCAAGCCCCTTTACCT<u>ATG</u>CTGCTGCTAACGCTGCTGCT
GCTGCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCT
CTTCTGCCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTC
CTGCCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGT
GGGGGATGGC<u>TAA</u>GAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGG
GCTGCCTTATTTAAAGTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGC
CCTGTTCATTAAGAAATTGTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGT
AAATATGTCTTTATAATAAACAGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 14

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56862

\<subunit 1 of 1, 73 aa, 1 stop

\<MW: 7879, pI: 7.21, NX(S/T): 0

MLLLTLLLLLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQ
PRGEGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 15

GGGACCCATGCGGCCGTGACCCCCGGCTCCCTAGAGGCCCAGCGCAGCCGCAGCGGACAAAG
GAGCATGTCCGCGCCGGGGAAGGCCCGTCCTCCGGCCGCCATAAGGCTCCGGTCGCCGCTGG
GCCCGCGCCGCGCTCCTGCCCGCCCGGGCTCCGGGGCGGCCCGCTAGGCCAGTGCGCCGCCG
CTCGCCCCGCAGGCCCCGGCCCGCAGC**ATG**GAGCCACCCGGACGCCGGCGGGGCCGCGCGCA
GCCGCCGCTGTTGCTGCCGCTCTCGCTGTTAGCGCTGCTCGCGCTGCTGGGAGGCGGCGGCG
GCGGCGGCGCCGCGGCGCTGCCCGCCGGCTGCAAGCACGATGGGCGGCCCCGAGGGGCTGGC
AGGGCGGCGGGCGCCGCCGAGGGCAAGGTGGTGTGCAGCAGCCTGGAACTCGCGCAGGTCCT
GCCCCCAGATACTCTGCCCAACCGCACGGTCACCCTGATTCTGAGTAACAATAAGATATCCG
AGCTGAAGAATGGCTCATTTTCTGGGTTAAGTCTCCTTGAAAGATTGGACCTCCGAAACAAT
CTTATTAGTAGTATAGATCCAGGTGCCTTCTGGGGACTGTCATCTCTAAAAAGATTGGATCT
GACAAACAATCGAATAGGATGTCTGAATGCAGACATATTTCGAGGACTCACCAATCTGGTTC
GGCTAAACCTTTCGGGGAATTTGTTTTCTTCATTATCTCAAGGAACTTTTGATTATCTTGCG
TCATTACGGTCTTTGGAATTCCAGACTGAGTATCTTTTGTGTGACTGTAACATACTGTGGAT
GCATCGCTGGGTAAAGGAGAAGAACATCACGGTACGGGATACCAGGTGTGTTTATCCTAAGT
CACTGCAGGCCCAACCAGTCACAGGCGTGAAGCAGGAGCTGTTGÀCATGCGACCÓTCCGCTT
GAATTGCCGTCTTTCTACATGACTCCATCTCATCGCCAAGTTGTGTTTGAAGGAGACAGCCT
TCCTTTCCAGTGCATGGCTTCATATATTGATCAGGACATGCAAGTGTTGTGGTATCAGGATG
GGAGAATAGTTGAAACCGATGAATCGCAAGGTATTTTTGTTGAAAAGAACATGATTCACAAC
TGCTCCTTGATTGCAAGTGCCCTAACCATTTCTAATATTCAGGCTGGATCTACTGGAAATTG
GGGCTGTCATGTCCAGACCAAACGTGGGAATAATACGAGGACTGTGGATATTGTGGTATTAG
AGAGTTCTGCACAGTACTGTCCTCCAGAGAGGGTGGTAAACAACAAAGGTGACTTCAGATGG
CCCAGAACATTGGCAGGCATTACTGCATATCTGCAGTGTACGCGGAACACCCATGGCAGTGG
GATATATCCCGGAAACCCACAGGATGAGAGAAAAGCTTGGCGCAGATGTGATAGAGGTGGCT
TTTGGGCAGATGATGATTATTCTCGCTGTCAGTATGCAAATGATGTCACTAGAGTTCTTTAT
ATGTTTAATCAGATGCCCCTCAATCTTACCAATGCCGTGGCAACAGCTCGACAGTTACTGGC
TTACACTGTGGAAGCAGCCAACTTTTCTGACAAAATGGATGTTATATTTGTGGCAGAAATGA
TTGAAAAATTTGGAAGATTTACCAAGGAGGAAAAATCAAAAGAGCTAGGTGACGTGATGGTT
GACATTGCAAGTAACATCATGTTGGCTGATGAACGTGTCCTGTGGCTGGCGCAGAGGGAAGC
TAAAGCCTGCAGTAGGATTGTGCAGTGTCTTCAGCGCATTGCTACCTACCGGCTAGCCGGTG
GAGCTCACGTTTATTCAACATATTCACCCAATATTGCTCTGGAAGCTTATGTCATCAAGTCT
ACTGGCTTCACGGGGATGACCTGTACCGTGTTCCAGAAAGTGGCAGCCTCTGATCGTACAGG
ACTTTCGGATTATGGGAGGCGGGATCCAGAGGGAAACCTGGATAAGCAGCTGAGCTTTAAGT
GCAATGTTTCAAATACATTTTCGAGTCTGGCACTAAAGGTATGTTACATTCTGCAATCATTT
AAGACTATTTACAGT**TAA**ATTAGAATGCTCCAAATGTTCTGCTTCGCAAAATAACCTTATTA
AAAGATTTTTTTTTGCAGGAAGATAGGTATTATTGCTTTTGCTACTGTTTTAAAGAAAACTA
ACCAGGAAGAACTGCATTACGACTTTCAAGGGCCCTAGGCATTTTTGCCTTTGATTCCCTTT
CTTCACATAAAAATATCAGAAATTACATTTTATAACTGCAGTGGTATAAATGCAAATATACT
ATTGTTACATGTGAAAAAATTTTATTTGACTTAAAAGTTTATTTATTTGTTTTTTTGCTCCT
GATTTTAAGACAATAAGATGTTTTCATGGGCCCCTAAAAGTATCATGAGCCTTTGGCACTGC
GCCTGCCAAGCCTAGTGGAGAAGTCAACCCTGAGACCAGGTGTTTAATCAAGCAAGCTGTAT
ATCAAAATTTTTGGCAGAAAACACAAATATGTCATATATCTTTTTTTTAAAAAAAAGTATTTCA
TTGAAGCAAGCAAAATGAAAGCATTTTTACTGATTTTTAAAATTGGTGCTTTAGATATATTT
GACTACACTGTATTGAAGCAAATAGAGGAGGCACAACTCCAGCACCCTAATGGAACCACATT
TTTTTCACTTAGCTTTCTGTGGGCATGTGTAATTGTATTCTCTGCGGTTTTTAATCTCACAG
TACTTTATTTCTGTCTTGTCCCTCAATAATATCACAAACAATATTCCAGTCATTTTAATGGC
TGCATAATAACTGATCCAACAGGTGTTAGGTGTTCTGGTTTAGTGTGAGCACTCAATAAATA
TTGAATGAATGAACGAAAAAAAAAAAAAAAAA

15 /270

FIGURE 16

MEPPGRRRGRAQPPLLLPLSLLALLALLGGGGGGGAAALPAGCKHDGRPRGAGRAAGAAEGK

VVCSSLELAQVLPPDTLPNRTVTLILSNNKISELKNGSFSGLSLLERLDLRNNLISSIDPGA

FWGLSSLKRLDLTNNRIGCLNADIFRGLTNLVRLNLSGNLFSSLSQGTFDYLASLRSLEFQT

EYLLCDCNILWMHRWVKEKNITVRDTRCVYPKSLQAQPVTGVKQELLTCDPPLELPSFYMTP

SHRQVVFEGDSLPFQCMASYIDQDMQVLWYQDGRIVETDESQGIFVEKNMIHNCSLIASALT

ISNIQAGSTGNWGCHVQTKRGNNTRTVDIVVLESSAQYCPPERVVNNKGDFRWPRTLAGITA

YLQCTRNTHGSGIYPGNPQDERKAWRRCDRGGFWADDDYSRCQYANDVTRVLYMFNQMPLNL

TNAVATARQLLAYTVEAANFSDKMDVIFVAEMIEKFGRFTKEEKSKELGDVMVDIASNIMLA

DERVLWLAQREAKACSRIVQCLQRIATYRLAGGAHVYSTYSPNIALEAYVIKSTGFTGMTCT

VFQKVAASDRTGLSDYGRRDPEGNLDKQLSFKCNVSNTFSSLALKVCYILQSFKTIYS

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 13-40 (type II)

N-glycosylation site.

amino acids 81-85, 98-102, 159-163, 206-210, 301-305, 332-336,
433-437, 453-457, 592-596

N-myristoylation site.

amino acids 29-35, 30-36, 31-37, 32-38, 33-39, 34-40, 51-57,
57-63, 99-105, 123-129, 142-148, 162-168, 317-323, 320-326,
384-390, 403-409, 554-560

FIGURE 17

GCGTGGGG<u>ATG</u>TCTAGGAGCTCGAAGGTGGTGCTGGGCCTCTCGGTGCTGCTGACGGCGGCC

ACAGTGGCCGGCGTACATGTGAAGCAGCAGTGGGACCAGCAGAGGCTTCGTGACGGAGTTAT

CAGAGACATTGAGAGGCAAATTCGGAAAAAAGAAAACATTCGTCTTTTGGGAGAACAGATTA

TTTTGACTGAGCAACTTGAAGCAGAAAGAGAGAAGATGTTATTGGCAAAAGGATCTCAAAAA

TCA<u>TGA</u>CTTGAATGTGAAATATCTGTTGGACAGACAACACGAGTTTGTGTGTGTGTGTTGAT

GGAGAGTAGCTTAGTAGTATCTTCATCTTTTTTTTTGGTCACTGTCCTTTTAAACTTGATCA

AATAAAGGACAGTGGGTCATATAAGTTACTGCTTTCAGGGTCCCTTATATCTGAATAAAGGA

GTGTGGGCAGACACTTTTTGGAAGAGTCTGTCTGGGTGATCCTGGTAGAAGCCCCATTAGGG

TCACTGTCCAGTGCTTAGGGTTGTTACTGAGAAGCACTGCCGAGCTTGTGAGAAGGAAGGGA

TGGATAGTAGCATCCACCTGAGTAGTCTGATCAGTCGGCATGATGACGAAGCCACGAGAACA

TCGACCTCAGAAGGACTGGAGGAAGGTGAAGTGGAGGGAGAGACGCTCCTGATCGTCGAATCC

FIGURE 18

MSRSSKVVLGLSVLLTAATVAGVHVKQQWDQQRLRDGVIRDIERQIRKKENIRLLGEQIILT
EQLEAEREKMLLAKGSQKS

Signal peptide:
amino acids 1-21

FIGURE 19

CTGTCGTCTTTGCTTCAGCCGCAGTCGCCACTGGCTGCCTGAGGTGCTCTTACAGCCTGTTC
CAAGTGTGGCTTAATCCGTCTCCACCACCAGATCTTTCTCCGTGGATTCCTCTGCTAAGACC
GCTGCC**ATG**CCAGTGACGGTAACCCGCACCACCATCACAACCACCACGACGTCATCTTCGGG
CCTGGGGTCCCCCATGATCGTGGGGTCCCCTCGGGCCCTGACACAGCCCCTGGGTCTCCTTCGC
CTGCTGCAGCTGGTGTCTACCTGCGTGGCCTTCTCGCTGGTGGCTAGCGTGGGCGCCTGGAC
GGGGTCCATGGGCAACTGGTCCATGTTCACCTGGTGCTTCTGCTTCTCCGTGACCCTGATCA
TCCTCATCGTGGAGCTGTGCGGGCTCCAGGCCCGCTTCCCCCTGTCTTGGCGCAACTTCCCC
ATCACCTTCGCCTGCTATGCGGCCCTCTTCTGCCTCTCGGCCTCCATCATCTACCCCACCAC
CTATGTCCAGTTCCTGTCCCACGGCCGTTCGCGGGACCACGCCATCGCCGCCACCTTCTTCT
CCTGCATCGCGTGTGTGGCTTACGCCACCGAAGTGGCCTGGACCCGGGCCCGGCCCGGCGAG
ATCACTGGCTATATGGCCACCGTACCCGGGCTGCTGAAGGTGCTGGAGACCTTCGTTGCCTG
CATCATCTTCGCGTTCATCAGCGACCCCAACCTGTACCAGCACCAGCCGGCCCTGGAGTGGT
GCGTGGCGGTGTACGCCATCTGCTTCATCCTAGCGGCCATCGCCATCCTGCTGAACCTGGGG
GAGTGCACCAACGTGCTACCCATCCCCTTCCCCAGCTTCCTGTCGGGGCTGGCCTTGCTGTC
TGTCCTCCTCTATGCCACCGCCCTTGTTCTCTGGCCCCTCTACCAGTTCGATGAGAAGTATG
GCGGCCAGCCTCGGCGCTCGAGAGATGTAAGCTGCAGCCGCAGCCATGCCTACTACGTGTGT
GCCTGGGACCGCCGACTGGCTGTGGCCATCCTGACGGCCATCAACCTACTGGCGTATGTGGC
TGACCTGGTGCACTCTGCCCACCTGGTTTTTGTCAAGGTC**TAA**GACTCTCCCAAGAGGCTCC
CGTTCCCTCTCCAACCTCTTTGTTCTTCTTGCCCGAGTTTTCTTTATGGAGTACTTCTTTCC
TCCGCCTTTCCTCTGTTTTCCTCTTCCTGTCTCCCCTCCCTCCCACCTTTTTCTTTCCTTCC
CAATTCCTTGCACTCTAACCAGTTCTTGGATGCATCTTCTTCCTTCCCTTTCCTCTTGCTGT
TTCCTTCCTGTGTTGTTTTGTTGCCCACATCCTGTTTTCACCCCTGAGCTGTTTCTCTTTTT
CTTTTCTTTCTTTTTTTTTTTTTTTTTTTAAGACGGATTCTCACTCTGTGGCCCAGGCTGGAG
TGCAGTGGTGCGATCTCAGCTCACTGCAACCCCCGCCTCCTGGGTTCAAGCGATTCTCCTCC
CCCAGCCTCCCAAGTAGCTGGGAGGACAGGTGTGAGCTGCCGCACCCAGCCTGTTTCTCTTT
TTCCACTCTTCTTTTTTCTCATCTCTTTTCTGGGTTGCCTGTCGGCTTTCTTATCTGCCTGT
TTTGCAAGCACCTTCTCCTGTGTCCTTGGGAGCCCTGAGACTTCTTTCTCTCCTTGCCTCCA
CCCACCTCCAAAGGTGCTGAGCTCACATCCACACCCCTTGCAGCCGTCCATGCCACAGCCCC
CCAAGGGGCCCCATTGCCAAAGCATGCCTGCCCACCCTCGCTGTGCCTTAGTCAGTGTGTAC
GTGTGTGTGTGTGTGTGTTTGGGGGGTGGGGGGTGGGTAGCTGGGGATTGGGCCCTCTTTCT
CCCAGTGGAGGAAGGTGTGCAGTGTACTTCCCCTTTAAATTAAAAAACATATATATATATAT
ATTTGGAGGTCAGTAATTTCCAATGGGCGGGAGGCATTAAGCACCGACCCTGGGTCCCTAGG
CCCCGCCTGGCACTCAGCCTTGCCAGAGATTGGCTCCAGAATTTTTGCCAGGCTTACAGAACAC
CCACTGCCTAGAGGCCATCTTAAAGGAAGCAGGGGCTGGATGCCTTTCATCCCAACTATTCT
CTGTGGTATGAAAAAG

19/270

FIGURE 20

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58727
<subunit 1 of 1, 322 aa, 1 stop
<MW: 35274, pI: 8.57, NX(S/T): 1
MPVTVTRTTITTTTTSSSGLGSPMIVGSPRALTQPLGLLRLLQLVSTCVAFSLVASVGAWTG
SMGNWSMFTWCFCFSVTLIILIVELCGLQARFPLSWRNFPITFACYAALFCLSASIIYPTTY
VQFLSHGRSRDHAIAATFFSCIACVAYATEVAWTRARPGEITGYMATVPGLLKVLETFVACI
IFAFISDPNLYQHQPALEWCVAVYAICFILAAIAILLNLGECTNVLPIPFPSFLSGLALLSV
LLYATALVLWPLYQFDEKYGGQPRRSRDVSCSRSHAYYVCAWDRRLAVAILTAINLLAYVAD
LVHSAHLVFVKV
```

Important features:

Transmembrane domains:

amino acids 41-60 (type II), 66-85, 101-120, 137-153, 171-192, 205-226, 235-255 and 294-312

N-glycosylation site.

amino acids 66-69

Glycosaminoglycan attachment site.

amino acids 18-21

FIGURE 21

GAACGTGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTC
TTGAACTCGTGACCTCATGATCCGCTCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGACGC
CTGGCCAGCCTATGCATTTTTAAGAAATTATTCTGTATTAGGTGCTGTGCTAAACATTGGGCACTACAGTGACCA
AAACAGACTGAATTCCCCAAGAGCCAAAGACCAGTGAGGGAGACCAACAAGAAACAGGAAATGCAAAAGAGACCA
TTATTACTCACTATGACTAAGGGTCACAAATGGGGTACGTTGATGGAGAGTGATTTGTTAAGAGACTACAGAGGG
AGGACAGACTACCAAGAGGGGGGGCCAGGAAAGCTCCTCTGACGAGGTGGTATTTCAGCCCAAACTGGAAGAATGA
GAAAGAGCTAGCCAGCCATCAGAATAGTCCAGAAGAGATGGGGAGCACTACACTCACTACACTTTGGCCTGAGAA
AATAGCATGGGATTGGAGGAGGCTGGGGGAACACCACTTCTGCCGACCTGGGCAGGAGGCATTGAGGGCTTGAGA
AAGGGCAATGGCAGTAGCAGTAGAAAGGACAGGGTAGGAGCAGGGACTTTGCAGGTGGAATCATTAGGTCTTATC
AACAGATATGGGCAAGCAAAGCCAGGGGAGAATTGATGGTAATGCTGAGGTTTGGAGCCAGGCTAGATGGGACAG
TGGTGGGTGATGCAAAGGAAAGAGGTCAGGAAGCAGGGCCAGACGTGGGGAGAAGGTGTGGGGGTTTGGTTTCCA
TCTTGCCGAGTCTGCCGGAATGTGGATGGGAAGACCAAGAGGAGGAGCAAGGGGCAGAGGGGAAGGGAATCTTAA
AGAAGTCCTGGATGCCACACTCTTCTTCCTTCCTCCTCTTCCCTCTCCTCAGAGGTCTCACTCGTGGTTCTTCAT
TTCCTGCCCTGCCTCCATCTCCTCTGGGTGCTGGGAAAGTGGAGGATTAGCTGAAGTTTTGCTTCTCGGGGCCTG
TCTGAATCTCCATTGCTTTCTGGGAGGACATAATTCACCTGTCCTAGCTTCTTATCATCTTACATTTCCCTGTAG
CCACTGGGACATATGTGGTGTTCCTTCCTAGCTCCTGTCTCCTCCTCATGCCTTTGCTGGGTATGGGCATGTTAG
GGGGAAGGTCATTGCTGTCAGAGGGGCACTGACTTTCTAATGGTGTTACCCAAGGTGAATGTTGGAGACACAGTC
GCGATGCTGCCCAAGTCCCGGCGAGCCCTAACTATCCAGGAGATCGCTGCGCTGGCCAGGTCCTCCCTGCATGGT
ATGCAGCCCCTCCC**ATG**TTTCTGGCCACTTTGTCCTTTCTCCTCCCGTTTGCACATCCCTTTGGAACTGTTTCCT
GTGAGTACATGCTGGGGTCTCCCCTTTCTTCCCTTGCTCAGGTGAATCTCAGCCCCTTCTCCCACCCAAAGGTTC
ACATGGATCCTAACTACTGCCACCCTTCCACCTCCCTGCACCTGTGCTCCCTGGCCTGGTCCTTTACCAGGCTTC
TCCACCCTCCCCTATCTCCAGGTATTTCCCAGGTGGTGAAGGACCACGTGACCAAGCCTACCGCCATGGCCCAGG
GCCGAGTGGCTCACCTCATTGAGTGGAAGGGCTGGAGCAAGCCGAGTGACTCACCTGCTGCCCTGGAATCAGCCT
TTTCCTCCTATTCAGACCTCAGCGAGGGCGAACAAGAGGCTCGCTTTGCAGCAGGAGTGGCTGAGCAGTTTGCCA
TCGCGGAAGCCAAGCTCCGAGCATGGTCTTCGGTGGATGGCGAGGACTCCACTGATGACTCCTATGATGAGGACT
TTGCTGGGGGAATGGACACAGACATGGCTGGGCAGCTGCCCCTGGGGCCGCACCTCCAGGACCTGTTCACCGGCC
ACCGGTTCTCCCGGCCTGTGCGCCAGGGCTCCGTGGAGCCTGAGAGCGACTGCTCACAGACCGTGTCCCCAGACA
CCCCTGTGCTCTAGTCTGTGCAGCCTGGAGGATGGGTTGTTGGGCTCCCCGGCCCGGCTGGCCTCCCAGCTGCTGG
GCGATGAGCTGCTTCTCGCCAAACTGCCCCCCCAGCCGGGAAAGTGCCTTCCGCAGCCTGGGCCCACTGGAGGCCC
AGGACTCACTCTACAACTCGCCCCTCACAGAGTCCTGCCTTTCCCCCGCGGAGGAGGAGCCAGCCCCCTGCAAGG
ACTGCCAGCCACTCTGCCCACCACTAACGGGCAGCTGGGAACGGCAGCGGCAAGCCTCTGACCTGGCCTCTTCTG
GGGTGGTGTCCTTAGATGAGGATGAGGCAGAGCCAGAGGAACAG**TGA**CCCACATCATGCCTGGCAGTGGCATGCA
TCCCCCCGGCTGCTGCCAGGGGCAGAGCCTCTGTGCCCAAGTGTGGGCTCAAGGCTCCCAGCAGAGCTCCACAGCC
TAGAGGGCTCCTGGGAGCGCTCGCTTCTCCGTTGTGTGTTTTGCATGAAAGTGTTTGGAGAGGAGGCAGGGGCTG
GGCTGGGGGCGCATGTCCTGCCCCCACTCCCGGGGCTTGCCGGGGGTTGCCCGGGGCCTCTGGGGCATGGCTACA
GCTGTGGCAGACAGTGATGTTCATGTTCTTAAAAATGCCACACACACATTTCCTCCTCGGATAATGTGAACCACTA
AGGGGGGTTGTGACTGGGCTGTGTGAGGGTGGGGTGGGAGGGGGGCCCAGCAACCCCCCACCCTCCCCATGCCTCTC
TCTTCTCTGCTTTTCTTCTCACTTCCGAGTCCATGTGCAGTGCTTGATAGAATCACCCCCACCTGGAGGGGCTGG
CTCCTGCCCTCCCGGAGCCTATGGGTTGAGCCGTCCCTCAAGGGCCCCTGCCCAGCTGGGCTCGTGCTGTGCTTC
ATTCACCTCTCCATCGTCTCTAAATCTTCCTCTTTTTTCCTAAAGACAGAAGGTTTTTGGTCTGTTTTTTCAGTC
GGATCTTCTCTTCTCTGGGAGGCTTTGGAATGATGAAAGCATGTACCCTCCACCCTTTTCCTGGCCCCCTAATGG
GGCCTGGGCCCTTTCCCAACCCCTCCTAGGATGTGCGGGCAGTGTGCTGGCGCCTCCACAGCCAGCCGGGCTGCCC
ATTCACGCAGAGCTCTCTGAGCGGGAGGTGGAAGAAAGGATGGCTCTGGTTGCCACAGAGCTGGGACTTCATGTT
CTTCTAGAGAGGGCCACAAGAGGGCCACAGGGGTGGCCGGGAGTTGTCAGCTGATGCCTGCTGAGAGGCAGGAAT
TGTGCCAGTGAGTGACAGTCATGAGGGAGTGTCTCTTCTTGGGGAGGAAAGAAGGTAGAGCCTTTCTGTCTGAAT
ATCCTGCATTCCAAGGTCACTGGACTGTACGTTTTTATGGTTGTGGGAAGGGTGGGTGGCTTTAGAATTAAGGGC
CTTGTAGGCTTTGGCAGGTAAGAGGGCCCAAGGTAAGAACGAGAGCCAACGGGCACAAGCATTCTATATATAAGT
GGCTCATTAGGTGTTTATTTTGTTCTATTTAAGAATTTGTTTTATTAAATTAATATAAAAATCTTTGTAAATCTC
TAAAA

FIGURE 22

MFLATLSFLLPFAHPFGTVSCEYMLGSPLSSLAQVNLSPFSHPKVHMDPNYCHPSTSLHLCS

LAWSFTRLLHPPLSPGISQVVKDHVTKPTAMAQGRVAHLIEWKGWSKPSDSPAALESAFSSY

SDLSEGEQEARFAAGVAEQFAIAEAKLRAWSSVDGEDSTDDSYDEDFAGGMDTDMAGQLPLG

PHLQDLFTGHRFSRPVRQGSVEPESDCSQTVSPDTLCSSLCSLEDGLLGSPARLASQLLGDE

LLLAKLPPSRESAFRSLGPLEAQDSLYNSPLTESCLSPAEEEPAPCKDCQPLCPPLTGSWER

QRQASDLASSGVVSLDEDEAEPEEQ

Signal peptide:
amino acids 1-15

Casein kinase II phosphorylation site.
amino acids 123-127, 128-132, 155-159, 162-166, 166-170, 228-232, 285-289, 324-328

Tyrosine kinase phosphorylation site.
amino acids 44-52

N-myristoylation site.
amino acids 17-23, 26-32, 173-179

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 11-22

FIGURE 23

```
GGTTCCTGGGCGCTCTGTTACACAAGCAAGATACAGCCAGCCCCACCTAATTTTGTTTCCCT
GGCACCCTCCTGCTCAGTGCGACATTGTCACACTTAACCCATCTGTTTTCTCTAATGCACGA
CAGATTCCTTTCAGACAGGACAACTGTGATATTTCAGTTCCTGATTGTAAATACCTCCTAAG
CCTGAAGCTTCTGTTACTAGCCATTGTGAGCTTCAGTTTCTTCATCTGCAAAATGGGCATAA
TACAATCTATTCTTGCCACATCAAGGGATTGTTATTCCTTTAAAAAAAAACCAATACCAAAG
AAGCCTACAATGTTGGCCTTAGCCAAAATTCTGTTGATTTCAACGTTGTTTTATTCACTTCT
ATCGGGGAGCCATGGAAAAGAAAATCAAGACATAAACACAACACAGAACATTGCAGAAGTTT
TTAAAACAATGGAAAATAAACCTATTTCTTTGGAAAGTGAAGCAAACTTAAACTCAGATAAA
GAAAATATAACCACCTCAAATCTCAAGGCGAGTCATTCCCCTCCTTTGAATCTACCCAACAA
CAGCCACGGAATAACAGATTTCTCCAGTAACTCATCAGCAGAGCATTCTTTGGGCAGTCTAA
AACCCACATCTACCATTTCCACAAGCCCTCCCTTGATCCATAGCTTTGTTTCTAAAGTGCCT
TGGAATGCACCTATAGCAGATGAAGATCTTTTGCCCATCTCAGCACATCCCAATGCTACACC
TGCTCTGTCTTCAGAAAACTTCACTTGGTCTTTGGTCAATGACACCGTGAAAACTCCTGATA
ACAGTTCCATTACAGTTAGCATCCTCTCTTCAGAACCAACTTCTCCATCTGTGACCCCCTTG
ATAGTGGAACCAAGTGGATGGCTTACCACAAACAGTGATAGCTTCACTGGGTTTACCCCTTA
TCAAGAAAAAACAACTCTACAGCCTACCTTAAAATTCACCAATAATTCAAAACTCTTTCCAA
ATACGTCAGATCCCCAAAAAGAAAATAGAAATACAGGAATAGTATTCGGGGCCATTTTAGGT
GCTATTCTGGGTGTCTCATTGCTTACTCTTGTGGGCTACTTGTTGTGTGGAAAAAGGAAAAC
GGATTCATTTTCCCATCGGCGACTTATGACGACAGAAATGAACCAGTTCTGCGATTAGACA
ATGCACCGGAACCTTATGATGTGAGTTTTGGGAATTCTAGCTACTACAATCCAACTTTGAAT
GATTCAGCCATGCCAGAAAGTGAAGAAAATGCACGTGATGGCATTCCTATGGATGACATACC
TCCACTTCGTACTTCTGTATAGAACTAACAGCAAAAAGGCGTTAAACAGCAAGTGTCATCTA
CATCCTAGCCTTTTGACAAATTCATCTTTCAAAAGGTTACACAAAATTACTGTCACGTGGAT
TTTGTCAAGGAGAATCATAAAAGCAGGAGACCAGTAGCAGAAATGTAGACAGGATGTATCAT
CCAAAGGTTTTCTTTCTTACAATTTTTGGCCATCCTGAGGCATTTACTAAGTAGCCTTAATT
TGTATTTTAGTAGTATTTTCTTAGTAGAAAATATTTGTGGAATCAGATAAAACTAAAAGATT
TCACCATTACAGCCCTGCCTCATAACTAAATAATAAAAATTATTCCACCAAAAAATTCTAAA
ACAATGAAGATGACTCTTTACTGCTCTGCCTGAAGCCCTAGTACCATAATTCAAGATTGCAT
TTTCTTAAATGAAAATTGAAAGGGTGCTTTTTAAAGAAAATTTGACTTAAAGCTAAAAAGAG
GACATAGCCCAGAGTTTCTGTTATTGGGAAATTGAGGCAATAGAAATGACAGACCTGTATTC
TAGTACGTTATAATTTTCTAGATCAGCACACACATGATCAGCCCACTGAGTTATGAAGCTGA
CAATGACTGCATTCAACGGGGCCATGGCAGGAAAGCTGACCCTACCCAGGAAAGTAATAGCT
TCTTTAAAAGTCTTCAAAGGTTTTGGGAATTTTAACTTGTCTTAATATATCTTAGGCTTCAA
TTATTTGGGTGCCTTAAAAACTCAATGAGAATCATGGT
```

23./270

FIGURE 24

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58732
><subunit 1 of 1, 334 aa, 1 stop
><MW: 36294, pI: 4.98, NX(S/T): 13
MLALAKILLISTLFYSLLSGSHGKENQDINTTQNIAEVFKTMENKPISLESEANLNSDKENI
TTSNLKASHSPPLNLPNNSHGITDFSSNSSAEHSLGSLKPTSTISTSPPLIHSFVSKVPWNA
PIADEDLLPISAHPNATPALSSENFTWSLVNDTVKTPDNSSITVSILSSEPTSPSVTPLIVE
PSGWLTTNSDSFTGFTPYQEKTTLQPTLKFTNNSKLFPNTSDPQKENRNTGIVFGAILGAIL
GVSLLTLVGYLLCGKRKTDSFSHRRLYDDRNEPVLRLDNAPEPYDVSFGNSSYYNPTLNDSA
MPESEENARDGIPMDDIPPLRTSV
```

Signal peptide:
amino acids 1-23

Transmembrane domain:
amino acids 235-262

N-glycosylation site.
amino acids 30-34, 61-65, 79-83, 90-94, 148-152, 155-159,
163-167, 218-222, 225-229, 298-302, 307-311

FIGURE 25

AACAGGATCTCCTCTTGCAGTCTGCAGCCCAGGACGCTGATTCCAGCAGCGCCTTACCGCGC

AGCCCGAAGATTCACTATGGTGAAAATCGCCTTCAATACCCCTACCGCCGTGCAAAAGGAGG

AGGCGCGGCAAGACGTGGAGGCCCTCCTGAGCCGCACGGTCAGAACTCAGATACTGACCGGC

AAGGAGCTCCGAGTTGCCACCCAGGAAAAAGAGGGCTCCTCTGGGAGATGTATGCTTACTCT

CTTAGGCCTTTCATTCATCTTGGCAGGACTTATTGTTGGTGGAGCCTGCATTTACAAGTACT

TCATGCCCAAGAGCACCATTTACCGTGGAGAGATGTGCTTTTTTGATTCTGAGGATCCTGCA

AATTCCCTTCGTGGAGGAGAGCCTAACTTCCTGCCTGTGACTGAGGAGGCTGACATTCGTGA

GGATGACAACATTGCAATCATTGATGTGCCTGTCCCCAGTTTCTCTGATAGTGACCCTGCAG

CAATTATTCATGACTTTGAAAAGGGAATGACTGCTTACCTGGACTTGTTGCTGGGGAACTGC

TATCTGATGCCCCTCAATACTTCTATTGTTATGCCTCCAAAAAATCTGGTAGAGCTCTTTGG

CAAACTGGCGAGTGGCAGATATCTGCCTCAAACTTATGTGGTTCGAGAAGACCTAGTTGCTG

TGGAGGAAATTCGTGATGTTAGTAACCTTGGCATCTTTATTTACCAACTTTGCAATAACAGA

AAGTCCTTCCGCCTTCGTCGCAGAGACCTCTTGCTGGGTTTCAACAAACGTGCCATTGATAA

ATGCTGGAAGATTAGACACTTCCCCAACGAATTTATTGTTGAGACCAAGATCTGTCAAGAGT

AAGAGGCAACAGATAGAGTGTCCTTGGTAATAAGAAGTCAGAGATTTACAATATGACTTTAA

CATTAAGGTTTATGGGATACTCAAGATATTTACTCATGCATTTACTCTATTGCTTATGCTTT

AAAAAAAGGAAAAAAAAAAAAACTACTAACCACTGCAAGCTCTTGTCAAATTTTAGTTTAAT

TGGCATTGCTTGTTTTTTGAAACTGAAATTACATGAGTTTCATTTTTTCTTTGCATTTATAG

GGTTTAGATTTCTGAAAGCAGCATGAATATATCACCTAACATCCTGACAATAAATTCCATCC

GTTGTTTTTTTTGTTTGTTTGTTTTTTCTTTTCCTTTAAGTAAGCTCTTTATTCATCTTATG

GTGGAGCAATTTTAAAATTTGAAATATTTTAAATTGTTTTTGAACTTTTTGTGTAAAATATA

TCAGATCTCAACATTGTTGGTTTCTTTTGTTTTTCATTTTGTACAACTTTCTTGAATTTAGA

AATTACATCTTTGCAGTTCTGTTAGGTGCTCTGTAATTAACCTGACTTATATGTGAACAATT

TTCATGAGACAGTCATTTTTAACTAATGCAGTGATTCTTTCTCACTACTATCTGTATTGTGG

AATGCACAAAATTGTGTAGGTGCTGAATGCTGTAAGGAGTTTAGGTTGTATGAATTCTACAA

CCCTATAATAAATTTTACTCTATACAAAAAAAAAAAAAAAAAAAA

FIGURE 26

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58828

<subunit 1 of 1, 263 aa, 1 stop

<MW: 29741, pI: 5.74, NX(S/T): 1

MVKIAFNTPTAVQKEEARQDVEALLSRTVRTQILTGKELRVATQEKEGSSGRCMLTLLGLSF
ILAGLIVGGACIYKYFMPKSTIYRGEMCFFDSEDPANSLRGGEPNFLPVTEEADIREDDNIA
IIDVPVPSFSDSDPAAIIHDFEKGMTAYLDLLLGNCYLMPLNTSIVMPPKNLVELFGKLASG
RYLPQTYVVREDLVAVEEIRDVSNLGIFIYQLCNNRKSFRLRRRDLLLGFNKRAIDKCWKIR
HFPNEFIVETKICQE

Type II transmembrane domain:
amino acids 53-75

N-glycosylation site.
amino acids 166-170

Casein kinase II phosphorylation site.
amino acids 35-39, 132-136, 134-138

N-myristoylation site.
amino acids 66-72, 103-109

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 63-74

FIGURE 27

GGAGGAGGGAGGGCGGGCAGGCGCCAGCCCAGAGCAGCCCCGGGCACCAGCACGGACTCTCT

CTTCCAGCCCAGGTGCCCCCCACTCTCGCTCCATTCGGCGGGAGCACCCAGTCCTGTACGCC

AAGGAACTGGTCCTGGGGGCACC**ATG**GTTTCGGCGGCAGCCCCCAGCCTCCTCATCCTTCTG

TTGCTGCTGCTGGGGTCTGTGCCTGCTACCGACGCCCGCTCTGTGCCCCTGAAGGCCACGTT

CCTGGAGGATGTGGCGGGTAGTGGGGAGGCCGAGGGCTCGTCGGCCTCCTCCCCGAGCCTCC

CGCCACCCTGGACCCCGGCCCTCAGCCCCACATCGATGGGGCCCCAGCCCACAACCCTGGGG

GGCCCATCACCCCCCACCAACTTCCTGGATGGGATAGTGGACTTCTTCCGCCAGTACGTGAT

GCTGATTGCTGTGGTGGGCTCCCTGGCCTTTCTGCTGATGTTCATCGTCTGTGCCGCGGTCA

TCACCCGGCAGAAGCAGAAGGCCTCGGCCTATTACCCATCGTCCTTCCCCAAGAAGAAGTAC

GTGGACCAGAGTGACCGGGCCGGGGGCCCCCGGGCCTTCAGTGAGGTCCCCGACAGAGCCCC

CGACAGCAGGCCCGAGGAAGCCCTGGATTCCTCCCGGCAGCTCCAGGCCGACATCTTGGCCG

CCACCCAGAACCTCAAGTCCCCCACCAGGGCTGCACTGGGCGGTGGGGACGGAGCCAGGATG

GTGGAGGGCAGGGGCGCAGAGGAAGAGGAGAAGGGCAGCCAGGAGGGGGACCAGGAAGTCCA

GGGACATGGGGTCCCAGTGGAGACACCAGAGGCGCAGGAGGAGCCGTGCTCAGGGGTCCTTG

AGGGGGCTGTGGTGGCCGGTGAGGGCCAAGGGGAGCTGGAAGGGTCTCTCTTGTTAGCCCAG

GAAGCCCAGGGACCAGTGGGTCCCCCCGAAAGCCCCTGTGCTTGCAGCAGTGTCCACCCCAG

TGTC**TAA**CAGTCCTCCCGGGCTGCCAGCCCTGACTGTCGGGCCCCCAAGTGGTCACCTCCCC

GTGTATGAAAAGGCCTTCAGCCCTGACTGCTTCCTGACACTCCCTCCTTGGCCTCCCTGTGG

TGCCAATCCCAGCATGTGCTGATTCTACAGCAGGCAGAAATGCTGGTCCCCGGTGCCCCGGA

GGAATCTTACCAAGTGCCATCATCCTTCACCTCAGCAGCCCCAAAGGGCTACATCCTACAGC

ACAGCTCCCCTGACAAAGTGAGGGAGGGCACGTGTCCCTGTGACAGCCAGGATAAAACATCC

CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCCCGGCCCAAACTACTTTTTAAAACA

GCTACAGGGTAAAATCCTGCAGCACCCACTCTGGAAAATACTGCTCTTAATTTTCCTGAAGG

TGGCCCCCTGTTTCTAGTTGGTCCAGGATTAGGGATGTGGGGTATAGGGCATTTAAATCCTC

TCAAGCGCTCTCCAAGCACCCCCGGCCTGGGGGTGAGTTTCTCATCCCGCTACTGCTGCTGG

GATCAGGTTGAATGAATGGAACTCTTCCTGTCTGGCCTCCAAAGCAGCCTAGAAGCTGAGGG

GCTGTGTTTGAGGGGACCTCCACCCTGGGGAAGTCCGAGGGGCTGGGGAAGGGTTTCTGACG

CCCAGCCTGGAGCAGGGGGGCCCTGGCCACCCCCTGTTGCTCACACATTGTCTGGCAGCCTG

TGTCCACAATATTCGTCAGTCCTCGACAGGGAGCCTGGGCTCCGTCCTGCTTTAGGGAGGCT

CTGGCAGGAGGTCCTCTCCCCCATCCCTCCATCTGGGGCTCCCCCAACCTCTGCACAGCTCT

CCAGGTGCTGAGATATAATGCACCAGCACAATAAACCTTTATTCCGGCCTGAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA

27/270

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58852

><subunit 1 of 1, 283 aa, 1 stop

><MW: 29191, pI: 4.52, NX(S/T): 0

MVSAAAPSLLILLLLLLGSVPATDARSVPLKATFLEDVAGSGEAEGSSASSPSLPPPWTPAL

SPTSMGPQPTTLGGPSPPTNFLDGIVDFFRQYVMLIAVVGSLAFLLMFIVCAAVITRQKQKA

SAYYPSSFPKKKYVDQSDRAGGPRAFSEVPDRAPDSRPEEALDSSRQLQADILAATQNLKSP

TRAALGGGDGARMVEGRGAEEEEKGSQEGDQEVQGHGVPVETPEAQEEPCSGVLEGAVVAGE

GQGELEGSLLLAQEAQGPVGPPESPCACSSVHPSV

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 94-118

N-myristoylation site.

amino acids 18-24, 40-46, 46-52, 145-151, 192-198, 193-199, 211-217, 238-244, 242-248

FIGURE 29

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGA

GGGAGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCG

GCAAGGAGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAG**ATG**

AAGTTCCAGGGGCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGG

CCCCCTGCAGAGCGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCC

TGGGAGACGCCCTGAGCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCT

GGCTCTAAAGTCAGTGAGGCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAG

GCAGGTTCCAGGCTTTGGCGCAGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATG

CTCTGGGAAACACTGGGCACGAGATTGGCAGACAGGCAGAAGATGTCATTCGACACGGAGCA

GATGCTGTCCGCGGCTCCTGGCAGGGGGTGCCTGGCCACAGTGGTGCTTGGGAAACTTCTGG

AGGCCATGGCATCTTTGGCTCTCAAGGTGGCCTTGGAGGCCAGGGCCAGGGCAATCCTGGAG

GTCTGGGGACTCCGTGGGTCCACGGATACCCCGGAAACTCAGCAGGCAGCTTTGGAATGAAT

CCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGAGGGCCACCAAACTTTGGGACCAACAC

TCAGGGAGCTGTGGCCCAGCCTGGCTATGGTTCAGTGAGAGCCAGCAACCAGAATGAAGGGT

GCACGAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCAACTCTGGGGGAGGCAGCGGC

TCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAACAATGGCAGCAGCAGTGG

TGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAGTGGCGGCAGCAGTG

GTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGTCCTCCTGGGGA

TCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGACATAAACC

CGGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGGAGCGGGGAATCTGGGATTCAGGGCTTCA

GAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTGGA

GGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGT

TGGTGGAGTCAATACTGTGAACTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCT

GGAAGAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGA

AGCTCTCGCATCCCG**TGA**CCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACT

CCCTCCTTAAAACACCACCCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTA

GCTGCCCCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59212

><subunit 1 of 1, 440 aa, 1 stop

><MW: 42208, pI: 6.36, NX(S/T): 1

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
AGSKVSEALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
ADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGLGGQGQGNPGGLGTPWVHGYPGNSAGSFGM
NPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGS
GSQSGSSGSGSGSNGDNNNGSSSGGSSSGSSSGSSSGGSSGGSSGGSSGNSGGSRGDSGSESSW
GSSTGSSSGNHGGSGGGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRLL
GGSGDNYRGQGSSWGSGGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQ
RSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70,
74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158,
155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205,
218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252,
249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281,
279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298,
295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328,
323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387,
383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 31

GACCGGTCCCTCCGGTCCTGGATGTGCGGACTCTGCTGCAGCGAGGGCTGCAGGCCCGCCGGGCGGTGCTCACCG
TGCCCTGGCTGGTGGAGTTTCTCTCCTTTGCTGACCATGTTGTTCCCTTGCTGGAATATTACCGGGACATCTTCA
CTCTCCTGCTGCGCCTGCACCGGAGCTTGGTGTTGTCGCAGGAGAGTGAGGGGAAG**ATG**TGTTTCCTGAACAAGC
TGCTGCTACTTGCTGTCCTGGGCTGGCTTTTCCAGATTCCCACAGTCCCTGAGGACTTGTTCTTTCTGGAAGAGG
GTCCCTCATATGCCTTTGAGGTGGACACAGTAGCCCCAGAGCATGGCTTGGACAATGCGCCTGTGGTGGACCAGC
AGCTGCTCTACACCTGCTGCCCCTACATCGGAGAGCTCCGGAAACTGCTCGCTTCGTGGGTGTCAGGCAGTAGTG
GACGGAGTGGGGGCTTCATGAGGAAAATCACCCCCACCACTACCACCAGCCTGGGAGCCCAGCCTTCCCAGACCA
GCCAGGGGCTGCAGGCACAGCTCGCCCAGGCCTTTTTCCACAACCAGCCGCCCTCCTTGCGCCGGACCGTAGAGT
TCGTGGCAGAAAGAATTGGATCAAACTGTGTCAAACATATCAAGGCTACACTGGTGGCAGATCTGGTGCGCCAGG
CAGAGTCACTTCTCCAAGAGCAGCTGGTGACACAGGGAGAGGAAGGGGGAGACCCAGCCCAGCTGTTGGAGATCT
TGTGTTCCCAGCTGTGCCCTCACGGGGCCCAGGCATTGGCCCTGGGGCGGGAGTTCTGTCAAAGGAAGAGCCCTG
GGGCTGTGCGGGCGCTGCTTCCAGAGGAGACCCCGGCAGCCGTTCTGAGCAGTGCAGAGAACATTGCTGTGGGGC
TTGCAACAGAGAAAGCCTGTGCTTGGCTGTCAGCCAACATCACAGCACTGATCAGGAGGGAGGTGAAAGCAGCAG
TGAGTCGCACACTTCGAGCCCAGGGTCCTGAACCTGCTGCCCGGGGGGGAGCGGAGGGGCTGCTCCCGCGCC**TGAC**
GTGCTCTCCTTGGCCGTGGGGCCACGGGACCCTGACGAGGGAGTCTCCCCAGAGCATCTGGAACAGCTCCTAGGC
CAGCTGGGCCAGACGCTGCGGTGCCGCCAGTTCCTGTGCCCCACCTGCTGAGCAGCATCTGGCAAAGTGCTCTGTG
GAGTTAGCTTCCCTCCTCGTTGCAGATCAAATTCCTATCCTAGGGCCCCCGGCACAGTACAGGCTGGAGAGAGGG
CAGGCTCGAAGGCTTCTGCACATGCTGCTTTCCTTGTGGAAGGAAGACTTTCAGGGGCCGGTTCCGCTGCAGCTG
CTGCTGAGCCCAAGAAATGTGGGGCTTCTGGCAGACACAAGGCCAAGGGAGTGGGACTTGCTGCTATTCTTGCTA
CGGGAGCTGGTGGAGAAGGGTCTGATGGGACGGATGGAGATAGAGGCCTGCCTGGGCAGCCTCCACCAGGCCCAG
TGGCCAGGGGACTTTGCTGAAGAATTAGCAACACTGTCTAATCTGTTTCTAGCCGAGCCCCACCTGCCAGAACCC
CAGCTAAGAGCCTGTGAGTTGGTGCAGCCAAACCGGGGCACTGTGCTGGCCCAGAGCTAGGGCTGAGAAGTGGCC
CTGCCTTGGGCATTGCACCAGAACCCTGGACCCCCGCCTCACGAGGAGGCCCAAGTGCCCAATGCAGACCCTCAC
TGGTTGGGGTGTAGCTGGGTCTACAGTCAGACTTCCTGCTCTAAGGGTGTCACTGCCTGGCATCCCACCACGCGA
ATCCTAGAGGAAGGAGAGTTGGCCTGATTTGGGATTATGGCAGAAAAGTCCAGAGATGCCAGTCCTGGAGTAGAA
GAGGTGGTGTTTGTTTATCTCTTGGATACTAAATGAAATGAGGTGTGTGGGCTTGTCAACACAGAATTCAAGCCT
CATTTGCTATCCCAGCATCTCTTAAAACTTTGTAGTCTTGGAATTCATGACAGAGGCAAATGACTCCTGCTTAAC
TTATGAAGAAAGTTAAAACATGAATCTTGGGAGTCTACATTTTCTTATCACCAGGAGCTGGACTGCCATCTCCTT
ATAAATGCCTAACACAGGCCGGGTCTGGTGGCTCATGCCTGTAATCCCAGCACTTTGAGAGGCCTGAGGTCGGCG
GACTGCCTGAGGTCAGGAATTCAAGACCAGCCTGGCCAACATGGCAAAACCCCATCTCTACTAAAAATAAAAAAA
TTATTAGCTGGGCATGGTGGTGTGTGCCTGTAATCCCAGCTACTCAGGAGGATGAGGCAGGAGACCTGCTTGAAC
CTGGAGGTGGAGGTTGCAGTGAGCCGAGGTCGCACCACTGCACTCCAGTCTGGGTAACAGAGCGAGACTTTCTAG
AAAAAGCCTAACAAACAGATAAGGTAGGACTCAACCAACTGAAACCTGACTTTCCCCCTGTACCTTCAGCCCCTG
TGCAGGTAGTAACCTCTTGAGACCTCTCCCTGACCAGGGACCAAGCACAGGGCATTTAGAGCTTTTTAGAATAAA
CTGGTTTTCTTTAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
TTTTTTTTTTTTTTTTAAAAAGGGCTTTTATTAAAATTCTCCCCACACGATGGCTCCTGCAATCTGCCACAGCTC
TGGGGCGTGTCCTGTAGGGAAAGGCCCTGTTTTCCCTGAGGCGGGGCTGGGCTTGTCCATGGGTCCGCGGAGCTG
GCCGTGCTTGGCGCCCTGGCGTGTGTCTAGCTGCTTCTTGCCGGGCACAGAGCTGCGGGGTCTGGGGGCACCGGG
AGCTAAGAGCAGGCTCTGGTGCAGGGGTGGAGGCCTGTCTCTTAACCGACACCCTGAGGTGCTCCTGAGATGCTG
GGTCCACCCTGAGTGGCACGGGGAGCAGCTGTGGCCGGTGCTCCTTCYTAGGCCAGTCCTGGGGAAACTAAGCTC
GGGCCCTTCTTTGCAAAGACCGAGGATGGGGTGGGTGTGGGGGACTCATGGGGAATGGCCTGAGGAGCTACGTGT
GAAGAGGGCGCCGGTTTGTTGGCTGCAGCGGCCTGGAGCGCCTCTCTCCTGAGCCTCAGTTTCCCTTTCCGTCTA
ATGAAGAACATGCCGTCTCGGTGTCTCAGGGCTATTAGGACTTGCCCTCAGGAAGTGGCCTTGGACGAGCGTCAT
GTTATTTTCACAACTGTCCTGCGACGTTGGCCTGGGCACGTCATGGAATGGCCCATGTCCCTCTGCTGCGTGGAC
GTCGCGGTCGGGAGTGCGCAGCCAGAGGCGGGGCCAGACGTGCGCCTGGGGGTGAGGGGAGGCGCCCCGGGAGGG
CCTCACAGGAAGTTGGGCTCCCGCACCACCAGGCAGGGCGGGCTCCCGCCGCCGCCGCCGCCACCACCGTCCAGG
GGCCGGTAGACAAAGTGGAAGTCGCGCTTGGGCTCGCTGCGCAGCAGGTAGCCCTTGATGCAGTGCGGCAGCGCG
TCGTCCGCCAGCTGGAAGCAGCGCCCGTCCACCAGCACGAACAGCCGGTGCGCCT

FIGURE 32

MCFLNKLLLLAVLGWLFQIPTVPEDLFFLEEGPSYAFEVDTVAPEHGLDNAPVVDQQLLYTC
CPYIGELRKLLASWVSGSSGRSGGFMRKITPTTTTSLGAQPSQTSQGLQAQLAQAFFHNQPP
SLRRTVEFVAERIGSNCVKHIKATLVADLVRQAESLLQEQLVTQGEEGGDPAQLLEILCSQL
CPHGAQALALGREFCQRKSPGAVRALLPEETPAAVLSSAENIAVGLATEKACAWLSANITAL
IRREVKAAVSRTLRAQGPEPAARGERRGCSRA

Signal peptide:
amino acids 1-18

N-glycosylation site.
amino acids 244-248

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 89-93

Casein kinase II phosphorylation site.
amino acids 21-25, 167-171, 223-227

N-myristoylation site.
amino acids 100-106, 172-178, 207-213

Microbodies C-terminal targeting signal.
amino acids 278-282

FIGURE 33

```
TCCCTTGACAGGTCTGGTGGCTGGTTCGGGGTCTACTGAAGGCTGTCTTGATCAGGAAACTG
AAGACTCTCTGCTTTTGCCACAGCAGTTCCTGCAGCTTCCTTGAGGTGTGAACCCACATCCC
TGCCCCCAGGGCCACCTGCAGGACGCCGACACCTACCCCTCAGCAGACGCCGGAGAGAAATG
AGTAGCAACAAAGAGCAGCGGTCAGCAGTGTTCGTGATCCTCTTTGCCCTCATCACCATCCT
CATCCTCTACAGCTCCAACAGTGCCAATGAGGTCTTCCATTACGGCTCCCTGCGGGGCCGTA
GCCGCCGACCTGTCAACCTCAAGAAGTGGAGCATCACTGACGGCTATGTCCCCATTCTCGGC
AACAAGACACTGCCCTCTCGGTGCCACCAGTGTGTGATTGTCAGCAGCTCCAGCCACCTGCT
GGGCACCAAGCTGGGCCCTGAGATCGAGCGGGCTGAGTGTACAATCCGCATGAATGATGCAC
CCACCACTGGCTACTCAGCTGATGTGGGCAACAAGACCACCTACCGCGTCGTGGCCCATTCC
AGTGTGTTCCGCGTGCTGAGGAGGCCCCAGGAGTTTGTCAACCGGACCCCTGAAACCGTGTT
CATCTTCTGGGGGCCCCCGAGCAAGATGCAGAAGCCCCAGGGCAGCCTCGTGCGTGTGATCC
AGCGAGCGGGCCTGGTGTTCCCCAACATGGAAGCATATGCCGTCTCTCCCGGCCGCATGCGG
CAATTTGACGACCTCTTCCGGGGTGAGACGGGCAAGGACAGGGAGAAGTCTCATTCGTGGTT
GAGCACAGGCTGGTTTACCATGGTGATCGCGGTGGAGTTGTGTGACCACGTGCATGTCTATG
GCATGGTCCCCCCCAACTACTGCAGCCAGCGGCCCCGCCTCCAGCGCATGCCCTACCACTAC
TACGAGCCCAAGGGGCCGGACGAATGTGTCACCTACATCCAGAATGAGCACAGTCGCAAGGG
CAACCACCACCGCTTCATCACCGAGAAAAGGGTCTTCTCATCGTGGGCCCAGCTGTATGGCA
TCACCTTCTCCCACCCCTCCTGGACCTAGGCCACCCAGCCTGTGGGACCTCAGGAGGGTCAG
AGGAGAAGCAGCCTCCGCCCAGCCGCTAGGCCAGGGACCATCTTCTGGCCAATCAAGGCTTG
CTGGAGTGTCTCCCAGCCAATCAGGGCCTTGAGGAGGATGTATCCTCCAGCCAATCAGGGCC
TGGGGAATCTGTTGGCGAATCAGGGATTTGGGAGTCTATGTGGTTAATCAGGGGTGTCTTTC
TTGTGCAGTCAGGGTCTGCGCACAGTCAATCAGGGTAGAGGGGGTATTTCTGAGTCAATCTG
AGGCTAAGGACATGTCCTTTCCCATGAGGCCTTGGTTCAGAGCCCCAGGAATGGACCCCCCA
ATCACTCCCCACTCTGCTGGGATAATGGGGTCCTGTCCCAAGGAGCTGGGAACTTGGTGTTG
CCCCCTCAATTTCCAGCACCAGAAAGAGAGATTGTGTGGGGGTAGAAGCTGTCTGGAGGCCC
GGCCAGAGAATTTGTGGGGTTGTGGAGGTTGTGGGGGCGGTGGGGAGGTCCCAGAGGTGGGA
GGCTGGCATCCAGGTCTTGGCTCTGCCCTGAGACCTTGGACAAACCCTTCCCCCTCTCTGGG
CACCCTTCTGCCCACACCAGTTTCCAGTGCGGAGTCTGAGACCCTTTCCACCTCCCCTACAA
GTGCCCTCGGGTCTGTCCTCCCCGTCTGGACCCTCCCAGCCACTATCCCTTGCTGGAAGGCT
CAGCTCTTTGGGGGGTCTGGGGTGACCTCCCCACCTCCTGGAAAACTTTAGGGTATTTTTGC
GCAAACTCCTTCAGGGGTTGGGGGACTCTGAAGGAAACGGGACAAAACCTTAAGCTGTTTTCT
TAGCCCCTCAGCCAGCTGCCATTAGCTTGGCTCTTAAAGGGCCAGGCCTCCTTTTTCTGCCCT
CTAGCAGGGAGGTTTTCCAACTGTTGGAGGCGCCTTTGGGGCTGCCCCTTTGTCTGGAGTCA
CTGGGGGGCTTCCGAGGGTCTCCCTCGACCCTCTGTCGTCCTGGGATGGCTGTCGGGAGCTGT
ATCACCTGGGTTCTGTCCCCTGGCTCTGTATCAGGCACTTTATTAAAGCTGGGCCTCAGTGG
GGTGTGTTTGTCTCCTGCTCTTCTGGAGCCTGGAAGGAAAGGGCTTCAGGAGGAGGCTGTGA
GGCTGGAGGGACCAGATGGAGGAGGCCAGCAGCTAGCCATTGCACACTGGGGTGATGGGTGG
GGGCGGTGACTGCCCCAGACTTGGTTTTGTAATGATTTGTACAGGAATAAACACACCTACGC
TCCGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

33/270

FIGURE 34

MSSNKEQRSAVFVILFALITILILYSSNSANEVFHYGSLRGRSRRPVNLKKWSITDGYVPIL
GNKTLPSRCHQCVIVSSSSHLLGTKLGPEIERAECTIRMNDAPTTGYSADVGNKTTYRVVAH
SSVFRVLRRPQEFVNRTPETVFIFWGPPSKMQKPQGSLVRVIQRAGLVFPNMEAYAVSPGRM
RQFDDLFRGETGKDREKSHSWLSTGWFTMVIAVELCDHVHVYGMVPPNYCSQRPRLQRMPYH
YYEPKGPDECVTYIQNEHSRKGNHHRFITEKRVFSSWAQLYGITFSHPSWT

Signal peptide:
amino acids 1-29

Transmembrane domain:
amino acids 9-31 (type II)

N-glycosylation site.
amino acids 64-68, 115-119

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 50-54

Casein kinase II phosphorylation site.
amino acids 3-7, 29-33, 53-57, 197-201

Tyrosine kinase phosphorylation site.
amino acids 253-262

N-myristoylation site.
amino acids 37-43, 114-120, 290-294

FIGURE 35

```
GTTTCTCATAGTTGGCGTCTTCTAAAGGAAAAACACTAAAATGAGGAACTCAGCGGACCGGGAGCGACGCAGCTT
GAGGGAAGCATCCCTAGCTGTTGGCGCAGAGGGCGAGGCTGAAGCCGAGTGGCCCGAGGTGTCTGAGGGGCTGG
GGCAAAGGTGAAAGAGTTTCAGAACAAGCTTCCTGGAACCCATGACCCATGAAGTCTTGTCGACATTTATACCGT
CTGAGGGTAGCAGCTCGAAACTAGAAGAAGTGGAGTGTTGCCAGGGACGGCAGTATCTCTTTGTGTGACCCTGGC
GGCCTATGGGACGTTGGCTTCAGACCTTTGTGATACACCATGCTGCGTGGGACGATGACGGCGTGGAGAGGAATG
AGGCCTGAGGTCACACTGGCTTGCCTCCTCCTAGCCACAGCAGGCTGCTTTGCTGACTTGAACGAGGTCCCTCAG
GTCACCGTCCAGCCTGCGTCCACCGTCCAGAAGCCCGGAGGCACTGTGATCTTGGGCTGCGTGGTGGAACCTCCA
AGGATGAATGTAACCTGGCGCCTGAATGGAAAGGAGCTGAATGGCTCGGATGATGCTCTGGGTGTCCTCATCACC
CACGGGACCCTCGTCATCACTGCCCTTAACAACCACACTGTGGGACGGTACCAGTGTGTGGCCCGGATGCCTGCG
GGGGCTGTGGCCAGCGTGCCAGCCACTGTGACACTAGCCAATCTCCAGGACTTCAAGTTAGATGTGCAGCACGTG
ATTGAAGTGGATGAGGGAAACACAGCAGTCATTGCCTGCCACCTGCCTGAGAGCCACCCCAAAGCCCAGGTCCGG
TACAGCGTCAAACAAGAGTGGCTGGAGGCCTCCAGAGGTAACTACCTGATCATGCCCTCAGGGAACCTCCAGATT
GTGAATGCCAGCCAGGAGGACGAGGGCATGTACAAGTGTGCAGCCTACAACCCAGTGACCCAGGAAGTGAAAACC
TCCGGCTCCAGCGACAGGCTACGTGTGCGCCGCTCCACCGCTGAGGCTGCCCGCATCATCTACCCCCCAGAGGCC
CAAACCATCATCGTCACCAAAGGCCAGAGTCTCATTCTGGAGTGTGTGGCCAGTGGAATCCCACCCCCCACGGGTC
ACCTGGGCCAAGGATGGGTCCAGTGTCACCGGCTACAACAAGACGCGCTTCCTGCTGAGCAACCTCCTCATCGAC
ACCACCAGCGAGGAGGACTCAGGCACCTACCGCTGCATGGCCGACAATGGGGTTGGGCAGCCCGGGGCAGCGGTC
ATCCTCTACAATGTCCAGGTGTTTGAACCCCCTGAGGTCACCATGGAGCTATCCCAGCTGGTCATCCCCTGGGGC
CAGAGTGCCAAGCTTACCTGTGAGGTGCGTGGGAACCCCCCGCCCTCCGTGCTGTGGCTGAGGAATGCTGTGCCC
CTCATCTCCAGCCAGCGCCTCCGGCTCTCCCGCAGGGCCTGCGCGTGCTCAGCATGGGGCCTGAGGACGAAGGC
GTCTACCAGTGCATGGCCGAGAACGAGGTTGGGAGCGCCCATGCCGTAGTCCAGCTGCGGACCTCCAGGCCAAGC
ATAACCCCAAGGCTATGGCAGGATGCTGAGCTGGCTACTGGCACACCTCCTGTATCACCCTCCAAACTCGGCAAC
CCTGAGCAGATGCTGAGGGGGCAACCGGCGCTCCCCAGACCCCCAACGTCAGTGGGGCCTGCTTCCCCGAAGTGT
CCAGGAGAGAAGGGGCAGGGGGCTCCCGCCGAGGCTCCCATCATCCTCAGCTCGCCCCGCACCTCCAAGACAGAC
TCATATGAACTGGTGTGGCGGCCTCGGCATGAGGGCAGTGGCCGGGCGCCAATCCTCTACTATGTGGTGAAACAC
CGCAAGCAGGTCACAAATTCCTCTGACGATTGGACCATCTCTGGCATTCCAGCCAACCAGCACCGCCTGACCCTC
ACCAGACTTGACCCCGGGAGCTTGTATGAAGTGGAGATGGCAGCTTACAACTGTGCGGGAGAGGGGCAGACAGCC
ATGGTCACCTTCCGAACTGGACGGCGGCCCAAACCCGAGATCATGGCCAGCAAAGAGCAGCAGATCCAGAGAGAC
GACCCTGGAGCCAGTCCCCAGAGCAGCAGCCAGCCAGACCACGGCCGCCTCTCCCCCCCAGAAGCCCCGACAGG
CCCACCATCTCCACGGCCTCCGAGACCTCAGTGTACGTGACCTGGATTCCCGTGGGAATGGTGGGGTTCCCAATC
CAGTCCTTCCGTGTGGAGTACAAGAAGCTAAAGAAAGTGGGAGACTGGATTCTGGCCACCAGCGCCATCCCCCCA
TCGCGGCTGTCCGTGGAGATCACGGGCCTAGAGAAAGGCACCTCCTACAAGTTTCGAGTCCGGGCTCTGAACATG
CTGGGGGAGGAGAGCCGAGCCCAGCGCCCCCTCTCGGCCCTACGTGGTGTCGGGCTACAGCGGTCGCGTGTACGAGAGG
CCCGTGGCAGGTCCTTATATCACCTTCACGGATGCGGTCAATGAGACCACCATCATGCTCAAGTGGATGTACATC
CCAGCAAGTAACAACAACACCCCAATCCATGGCTTTTATATCTATTATCGACCCACAGACAGTGACAATGATAGT
GACTACAAGAAGGATATGGTGGAAGGGGGACAAGTACTGGCACTCCATCAGCCACCTGCAGCCAGAGACCTCCTAC
GACATTAAGATGCAGTGCTTCAATGAAGGAGGGGAGAGCGAGTTCAGCAACGTGATGATCTGTGAGACCAAAGCT
CGGAAGTCTTCTGGCCAGCCTGGTCGACTGCCACCCCCAACTCTGGCCCCACCACAGCCGCCCCTTCCTGAAACC
ATAGAGCGGCCGGTGGGCACTGGGGCCATGGTGGCTCGCTCCAGCGACCTGCCCTATCTGATTGTCGGGGTCGTC
CTGGGCTCCATCGTTCTCATCATCGTCACCTTCATCCCCTTCTGCTTGTGGAGGGCCTGGTCTAAGCAAAAACAT
ACAACAGACCTGGGTTTTCCTCGAAGTGCCCTTCCACCCTCCTGCCCGTATACTATGGTGCCATTGGGAGGACTC
CCAGGCCACCAGGCCAGTGGACAGCCCTACCTCAGTGGCATCAGTGGACGGGCCTGTGCTAATGGGATCCACATG
AATAGGGGCTGCCCCTCGGCTGCAGTGGGCTACCCGGGCATGAAGCCCCAGCAGCACTGCCCAGGCGAGCTTCAG
CAGCAGAGTGACACCAGCAGCCTGCTGAGGCAGACCCATCTTGGCAATGGATATGACCCCCAAAGTCACCAGATC
ACGAGGGGTCCCAAGTCTAGCCCGGACGAGGGCTCTTTCTTATACACACTGCCCGACGACTCCACTCACCAGCTG
CTGCAGCCCCATCACGACTGCTGCCAACGCCAGGAGCAGCCTGCTGCTGTGGGCCAGTCAGGGGTGAGGAGAGCC
CCCGACAGTCCTGTCCTGGAAGCAGTGTGGGACCCTCCATTTCACTCAGGGCCCCCATGCTGCTTGGGCCTTGTG
CCAGTTGAAGAGGTGGACAGTCCTGACTCCTGCCAAGTGAGTGGAGGAGACTGGTGTCCCCAGCACCCCGTAGGG
GCCTACGTAGGACAGGAACCTGGAATGCAGCTCTCCCCGGGGCCACTGGTCGCGTGTGTCTTTTGAAACACCACCT
CTCACAATTTAGGCAGAAGCTGATATCCCAGAAAGACTATATATTGTTTTTTTTTTTAAAAAAAAAAAGAAGAAAAA
AGAGACAGAGAAAATTGGTATTTATTTTTCTATTATAGCCATATTTATATATTTATGCACTTGTAAATAAATGTA
TATGTTTTATAATTCTGGAGAGACATAAGGAGTCCTACCCGTTGAGGTTGGAGAGGGAAAATAAAGAAGCTGCCA
CCTAACAGGAGTCACCCAGGAAAGCACCGCACAGGCTGGCGCGGGACAGACTCCTAACCTGGGGCCTCTGCAGTG
GCAGGCGAGGCTGCAGGAGGCCCACAGATAAGCTGGCAAGAGGAAGGATCCCAGGCACATGGTTCATCACGAGCA
TGAGGGAACAGCAAGGGGCACGGTATCACAGCCTGGAGACACCCACACAGATGGCTGGATCCGGTGCTACGGGAA
ACATTTTCCTAAGATGCCCATGAGAACAGACCAAGATGTGTACAGCACTATGAGCATTAAAAAAACCTTCCAGAAT
CAATAATCCGTGGCAACATATCTCTGTAAAAACAAACACTGTAACTTCTAAATAAATGTTTAGTCTTCCCTGTAAAA
```

FIGURE 36

MLRGTMTAWRGMRPEVTLACLLLATAGCFADLNEVPQVTVQPASTVQKPGGTVILGCVVEPP

RMNVTWRLNGKELNGSDDALGVLITHGTLVITALNNHTVGRYQCVARMPAGAVASVPATVTL

ANLQDFKLDVQHVIEVDEGNTAVIACHLPESHPKAQVRYSVKQEWLEASRGNYLIMPSGNLQ

IVNASQEDEGMYKCAAYNPVTQEVKTSGSSDRLRVRRSTAEAARIIYPPEAQTIIVTKGQSL

ILECVASGIPPPRVTWAKDGSSVTGYNKTRFLLSNLLIDTTSEEDSGTYRCMADNGVGQPGA

AVILYNVQVFEPPEVTMELSQLVIPWGQSAKLTCEVRGNPPPSVLWLRNAVPLISSQRLRLS

RRALRVLSMGPEDEGVYQCMAENEVGSAHAVVQLRTSRPSITPRLWQDAELATGTPPVSPSK

LGNPEQMLRGQPALPRPPTSVGPASPKCPGEKGQGAPAEAPIILSSPRTSKTDSYELVWRPR

HEGSGRAPILYYVVKHRKQVTNSSDDWTISGIPANQHRLTLTRLDPGSLYEVEMAAYNCAGE

GQTAMVTFRTGRRPKPEIMASKEQQIQRDDPGASPQSSSQPDHGRLSPPEAPDRPTISTASE

TSVYVTWIPRGNGGFPIQSFRVEYKKLKKVGDWILATSAIPPSRLSVEITGLEKGTSYKFRV

RALNMLGESEPSAPSRPYVVSGYSGRVYERPVAGPYITFTDAVNETTIMLKWMYIPASNNNT

PIHGFYIYYRPTDSDNDSDYKKDMVEGDKYWHSISHLQPETSYDIKMQCFNEGGESEFSNVM

ICETKARKSSGQPGRLPPPTLAPPQPPLPETIERPVGTGAMVARSSDLPYLIVGVVLGSIVL

IIVTFIPFCLWRAWSKQKHTTDLGFPRSALPPSCPYTMVPLGGLPGHQASGQPYLSGISGRA

CANGIHMNRGCPSAAVGYPGMKPQQHCPGELQQQSDTSSLLRQTHLGNGYDPQSHQITRGPK

SSPDEGSFLYTLPDDSTHQLLQPHHDCCQRQEQPAAVGQSGVRRAPDSPVLEAVWDPPFHSG

PPCCLGLVPVEEVDSPDSCQVSGGDWCPQHPVGAYVGQEPGMQLSPGPLVRVSFETPPLTI

Signal peptide:
amino acids 1-30

Transmembrane domain:
amino acids 16-30 (type II), 854-879

FIGURE 37

CGGGAGGCTGGGTCGTCATGATCCGGACCCCATTGTCGGCCTCTGCCCATCGCCTGCTCCTC

CCAGGCTCCCGCGGCCGACCCCCGCGCAAC**ATG**CAGCCCACGGGCCGCGAGGGTTCCCGCGC

GCTCAGCCGGCGGTATCTGCGGCGTCTGCTGCTCCTGCTACTGCTGCTGCTGCTGCGGCAGC

CCGTAACCCGCGCGGAGACCACGCCGGGCGCCCCAGAGCCCTCTCCACGCTGGGCTCCCCC

AGCCTCTTCACCACGCCGGGTGTCCCCAGCGCCCTCACTACCCCAGGCCTCACTACGCCAGG

CACCCCCAAAACCCTGGACCTTCGGGGTCGCGCGCAGGCCCTGATGCGGAGTTTCCCACTCG

TGGACGGCCACAATGACCTGCCCCAGGTCCTGAGACAGCGTTACAAGAATGTGCTTCAGGAT

GTTAACCTGCGAAATTTCAGCCATGGTCAGACCAGCCTGGACAGGCTTAGAGACGGCCTCGT

GGGTGCCCAGTTCTGGTCAGCCTCCGTCTCATGCCAGTCCCAGGACCAGACTGCCGTGCGCC

TCGCCCTGGAGCAGATTGACCTCATTCACCGCATGTGTGCCTCCTACTCTGAACTCGAGCTT

GTGACCTCAGCTGAAGGTCTGAACAGCTCTCAAAAGCTGGCCTGCCTCATTGGCGTGNAGGG

TGGTCACTCACTGGACAGCAGCCTCTCTGTGCTGCGCAGTTTCTATGTGCTGGGGGTGCGCT

ACCTGACACTTACCTTCACCTGCAGTACACCATGGGCAGAGAGTTCCACCAAGTTCAGACAC

CACATGTACACCAACGTCAGCGGATTGACAAGCTTTGGTGAGAAAGTAGTAGAGGAGTTGAA

CCGCCTGGGCATGATGATAGATTTGTCCTATGCATCGGACACCTTGATAAGAAGGGTCCTGG

AAGTGTCTCAGGCTCCTGTGATCTTCTCCCACTCAGCTGCCAGAGCTGTGTGTGACAATTTG

TTGAATGTTCCCGATGATATCCTGCAGCTTCTGAAGAACGGTGGCATCGTGATGGTGACACT

GTCCATGGGGGTGCTGCAGTGCAACCTGCTTGCTAACGTGTCCACTGTGGCAGATCACTTTG

ACCACATCAGGGCAGTCATTGGATCTGAGTTCATCGGGATTGGTGGAAATTATGACGGGACT

GGCCGGTTCCCTCAGGGGCTGGAGGATGTGTCCACATACCCAGTCCTGATAGAGGAGTTGCT

GAGTCGTASCTGGAGCGAGGAAGAGCTTCAAGGTGTCCTTCGTGGAAACCTGCTGCGGGTCT

TCAGACAAGTGGAAAAGGTGAGAGAGGAGAGCAGGGCGCAGAGCCCCGTGGAGGCTGAGTTT

CCATATGGGCAACTGAGCACATCCTGCCACTCCCACCTCGTGCCTCAGAATGGACACCAGGC

TACTCATCTGGAGGTGACCAAGCAGCCAACCAATCGGGTCCCCTGGAGGTCCTCAAATGCCT

CCCCATACCTTGTTCCAGGCCTTGTGGCTGCTGCCACCATCCCAACCTTCACCCAGTGGCTC

TGC**TGA**CACAGTCGGTCCCCGCAGAGGTCACTGTGGCAAAGCCTCACAAAGCCCCCTCTCCT

AGTTCATTCACAAGCATATGCTGAGAATAAACATGTTACACATGGAAAA

FIGURE 38

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59817

><subunit 1 of 1, 487 aa, 1 stop, 2 unknown

><MW: 53569.32, pI: 7.68, NX(S/T): 5

MQPTGREGSRALSRRYLRRLLLLLLLLLLLRQPVTRAETTPGAPRALSTLGSPSLFTTPGVPS
ALTTPGLTTPGTPKTLDLRGRAQALMRSFPLVDGHNDLPQVLRQRYKNVLQDVNLRNFSHGQ
TSLDRLRDGLVGAQFWSASVSCQSQDQTAVRLALEQIDLIHRMCASYSELELVTSAEGLNSS
QKLACLIGVXGGHSLDSSLSVLRSFYVLGVRYLTLTFTCSTPWAESSTKFRHHMYTNVSGLT
SFGEKVVEELNRLGMMIDLSYASDTLIRRVLEVSQAPVIFSHSAARAVCDNLLNVPDDILQL
LKNGGIVMVTLSMGVLQCNLLANVSTVADHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLEDV
STYPVLIEELLSRXWSEEELQGVLRGNLLRVFRQVEKVREESRAQSPVEAEFPYGQLSTSCH
SHLVPQNGHQATHLEVTKQPTNRVPWRSSNASPYLVPGLVAAATIPTFTQWLC

Important features of the protein:
Signal peptide:
amino acids 1-36

Transmembrane domain:
amino acids 313-331

N-glycosylation sites.
amino acids 119-122, 184-187, 243-246 and 333-336

N-myristoylation sites.
amino acids 41-46, 59-64, 73-78, 133-138, 182-187, 194-199, 324-329, 354-359, 357-362, 394-399, 427-432 and 472-477.

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 136-146

FIGURE 39

TGCTAGGCTCTGTCCCACAATGCACCCGAGAGCAGGAGCTGAAAGCCTCTAACACCCACAGA

TCCCTCTATGACTGCAATGTGAGGTGTCCGGCTTTGCTGGCCCAGCAAGCCTGATAAGCATG

AAGCTCTTATCTTTGGTGGCTGTGGTCGGGTGTTTGCTGGTGCCCCCAGCTGAAGCCAACAA

GAGTTCTGAAGATATCCGGTGCAAATGCATCTGTCCACCTTATAGAAACATCAGTGGGCACA

TTTACAACCAGAATGTATCCCAGAAGGACTGCAACTGCCTGCACGTGGTGGAGCCCATGCCA

GTGCCTGGCCATGACGTGGAGGCCTACTGCCTGCTGTGCGAGTGCAGGTACGAGGAGCGCAG

CACCACCACCATCAAGGTCATCATTGTCATCTACCTGTCCGTGGTGGGTGCCCTGTTGCTCT

ACATGGCCTTCCTGATGCTGGTGGACCCTCTGATCCGAAAGCCGGATGCATACACTGAGCAA

CTGCACAATGAGGAGGAGAATGAGGATGCTCGCTCTATGGCAGCAGCTGCTGCATCCCTCGG

GGGACCCCGAGCAAACACAGTCCTGGAGCGTGTGGAAGGTGCCCAGCAGCGGTGGAAGCTGC

AGGTGCAGGAGCAGCGGAAGACAGTCTTCGATCGGCACAAGATGCTCAGCTAGATGGGCTGG

TGTGGTTGGGTCAAGGCCCCAACACCATGGCTGCCAGCTTCCAGGCTGGACAAAGCAGGGGG

CTACTTCTCCCTTCCCTCGGTTCCAGTCTTCCCTTTAAAAGCCTGTGGCATTTTTCCTCCTT

CTCCCTAACTTTAGAAATGTTGTACTTGGCTATTTTGATTAGGGAAGAGGGATGTGGTCTCT

GATCTCTGTTGTCTTCTTGGGTCTTTGGGGTTGAAGGGAGGGGGAAGGCAGGCCAGAAGGGA

ATGGAGACATTCGAGGCGGCCTCAGGAGTGGATGCGATCTGTCTCTCCTGGCTCCACTCTTG

CCGCCTTCCAGCTCTGAGTCTTGGGAATGTTGTTACCCTTGGAAGATAAAGCTGGGTCTTCA

GGAACTCAGTGTCTGGGAGGAAAGCATGGCCCAGCATTCAGCATGTGTTCCTTTCTGCAGTG

GTTCTTATCACCACCTCCCTCCCAGCCCCGGCGCCTCAGCCCCAGCCCCAGCTCCAGCCCTG

AGGACAGCTCTGATGGGAGAGCTGGGCCCCCTGAGCCCACTGGGTCTTCAGGGTGCACTGGA

AGCTGGTGTTCGCTGTCCCCTGTGCACTTCTCGCACTGGGGCATGGAGTGCCCATGCATACT

CTGCTGCCGGTCCCCTCACCTGCACTTGAGGGGTCTGGGCAGTCCCTCCTCTCCCCAGTGTC

CACAGTCACTGAGCCAGACGGTCGGTTGGAACATGAGACTCGAGGCTGAGCGTGGATCTGAA

CACCACAGCCCCTGTACTTGGGTTGCCTCTTGTCCCTGAACTTCGTTGTACCAGTGCATGGA

GAGAAAATTTTGTCCTCTTGTCTTAGAGTTGTGTGTAAATCAAGGAAGCCATCATTAAATTG

TTTTATTTCTCTCA

FIGURE 40

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60278

<subunit 1 of 1, 183 aa, 1 stop

<MW: 20574, pI: 6.60, NX(S/T): 3

MKLLSLVAVVGCLLVPPAEANKSSEDIRCKCICPPYRNISGHIYNQNVSQKDCNCLHVVEPM

PVPGHDVEAYCLLCECRYEERSTTTIKVIIVIYLSVVGALLLYMAFLMLVDPLIRKPDAYTE

QLHNEEENEDARSMAAAAASLGGPRANTVLERVEGAQQRWKLQVQEQRKTVFDRHKMLS

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 90-112

N-glycosylation sites.

amino acids 21-24, 38-41 and 47-50

FIGURE 41

```
AGCGGGTCTCGCTTGGGTTCCGCTAATTTCTGTCCTGAGGCGTGAGACTGAGTTCATAGGGTCCTGGGTCCCCGA
ACCAGGAAGGGTTGAGGGAACACAATCTGCAAGCCCCCGCGACCCAAGTGAGGGGCCCCGTGTTGGGGTCCTCCC
TCCCTTTGCATTCCCACCCCTCCGGGCTTTGCGTCTTCCTGGGGACCCCCTCGCCGGGAGATGGCCGCGTTGATG
CGGAGCAAGGATTCGTCCTGCTGCCTGCTCCTACTGGCCGCGGTGCTGATGGTGGAGAGCTCACAGATCGGCAGT
TCGCGGGCCAAACTCAACTCCATCAAGTCCTCTCTGGGCGGGGAGACGCCTGGTCAGGCCGCCAATCGATCTGCG
GGCATGTACCAAGGACTGGCATTCGGCGGCAGTAAGAAGGGCAAAAACCTGGGGCAGGCCTACCCTTGTAGCAGT
GATAAGGAGTGTGAAGTTGGGAGGTATTGCCACAGTCCCCACCAAGGATCATCGGCCTGCATGGTGTGTCGGAGA
AAAAAGAAGCGCTGCCACCGAGATGGCATGTGCTGCCCCAGTACCCGCTGCAATAATGGCATCTGTATCCCAGTT
ACTGAAAGCATCTTAACCCCTCACATCCCGGCTCTGGATGGTACTCGGCACAGAGATCGAAACCACGGTCATTAC
TCAAACCATGACTTGGGATGGCAGAATCTAGGAAGACCACACACTAAGATGTCACATATAAAAGGGCATGAAGGA
GACCCCTGCCTACGATCATCAGACTGCATTGAAGGGTTTTGCTGTGCTCGTCATTTCTGGACCAAAATCTGCAAA
CCAGTGCTCCATCAGGGGGAAGTCTGTACCAAACAACGCAAGAAGGGTTCTCATGGGCTGGAAATTTTCCAGCGT
TGCGACTGTGCGAAGGGCCTGTCTTGCAAAGTATGGAAAGATGCCACCTACTCCTCCAAAGCCAGACTCCATGTG
TGTCAGAAAATTTGATCACCATTGAGGAACATCATCAATTGCAGACTGTGAAGTTGTGTATTTAATGCATTATAG
CATGGTGGAAAATAAGGTTCAGATGCAGAAGAATGGCTAAAATAAGAAACGTGATAAGAATATAGATGATCACAA
AAAGGGAGAAAGAAAACATGAACTGAATAGATTAGAATGGGTGACAAATGCAGTGCAGCCAGTGTTTCCATTATG
CAACTTGTCTATGTAAATAATGTACACATTTGTGGAAAATGCTATTATTAAGAGAACAAGCACACAGTGGAAATT
ACTGATGAGTAGCATGTGACTTTCCAAGAGTTTAGGTTGTGCTGGAGGAGAGGTTTCCTTCAGATTGCTGATTGC
TTATACAAATAACCTACATGCCAGATTTCTATTCAACGTTAGAGTTTAACAAAATACTCCTAGAATAACTTGTTA
TACAATAGGTTCTAAAAATAAAATTGCTAAACAAGAAATGAAAACATGGAGCATTGTTAATTTACAACAGAAAAT
TACCTTTTGATTTGTAACACTACTTCTGCTGTTCAATCAAGAGTCTTGGTAGATAAGAAAAAAAATCAGTCAATAT
TTCCAAATAATTGCAAAATAATGGCCAGTTGTTTAGGAAGGCCTTTAGGAAGACAAATAAATAACAAACAAACAG
CCACAAATACTTTTTTTTCAAAATTTTAGTTTTACCTGTAATTAATAAGAACTGATACAAGACAAAAACAGTTCC
TTCAGATTCTACGGAATGACAGTATATCTCTCTTTATCCTATGTGATTCCTGCTCTGAATGCATTATATTTTCCA
AACTATACCCATAAATTGTGACTAGTAAAATACTTACACAGAGCAGAATTTTCACAGATGGCAAAAAAATTTAAA
GATGTCCAATATATGTGGGAAAAGAGCTAACAGAGAGATCATTATTTCTTAAAGATTGGCCATAACCTATATTTT
GATAGAATTAGATTGGTAAATACATGTATTCATACATACTCTGTGGTAATAGAGACTTAAGCTGGATCTGTACTG
CACTGGAGTAAGCAAGAAAATTGGGAAAACTTTTTCGTTTGTTCAGGTTTTGGCAACACATAGATCATATGTCTG
AGGCACAAGTTGGCTGTTCATCTTTGAAACCAGGGGATGCACAGTCTAAATGAATATCTGCATGGGATTTGCTAT
CATAATATTTACTATGCAGATGAATTCAGTGTGAGGTCCTGTGTCCGTACTATCCTCAAATTATTTATTTTATAG
TGCTGAGATCCTCAAATAATCTCAATTTCAGGAGGTTTCACAAAATGTACTCCTGAAGTAGACAGAGTAGTGAGG
TTTCATTGCCCTCTATAAGCTTCTGACTAGCCAATGGCATCATCCAATTTTCTTCCCAAACCTCTGCAGCATCTG
CTTTATTGCCAAAGGGCTAGTTTCGGTTTTCTGCAGCCATTGCGGTTAAAAAAATATAAGTAGGATAACTTGTAAA
ACCTGCATATTGCTAATCTATAGACACCACAGTTTCTAAATTCTTTGAAACCACTTTACTACTTTTTTTTAAACTT
AACTCAGTTCTAAATACTTTGTCTGGAGCACAAAACAATAAAAGGTTATCTTATAGTCGTGACTTTAAACTTTTG
TAGACCACAATTCACTTTTTAGTTTTCTTTTACTTAAATCCCATCTGCAGTCTCAAATTTAAGTTCTCCCAGTAG
AGATTGAGTTTGAGCCTGTATATCTATTAAAAAATTTCAACTTCCCACATATATTTACTAAGATGATTAAGACTTA
CATTTTCTGCACAGGTCTGCAAAAACAAAAATTATAAACTAGTCCATCCAAGAACCAAAGTTTGTATAAACAGGT
TGCTATAAGCTTGTGAAATGAAAATGGAACATTTCAATCAAACATTTCCTATATAACAATTATTATATTTACAAT
TTGGTTTCTGCAATATTTTTCTTATGTCCACCCTTTTAAAAATTATTATTTGAAGTAATTTATTTTACAGGAAATG
TTAATGAGATGTATTTTCTTATAGAGATATTTCTTACAGAAAGCTTTGTAGCAGAATATATTTGCAGCTATTGAC
TTTGTAATTTAGGAAAAATGTATAATAAGATAAAATCTATTAAATTTTTTCTCCTCTAAAAACTGAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA
```

411270

FIGURE 42

MAALMRSKDSSCCLLLLAAVLMVESSQIGSSRAKLNSIKSSLGGETPGQAANRSAGMYQGLA

FGGSKKGKNLGQAYPCSSDKECEVGRYCHSPHQGSSACMVCRRKKKRCHRDGMCCPSTRCNN

GICIPVTESILTPHIPALDGTRHRDRNHGHYSNHDLGWQNLGRPHTKMSHIKGHEGDPCLRS

SDCIEGFCCARHFWTKICKPVLHQGEVCTKQRKKGSHGLEIFQRCDCAKGLSCKVWKDATYS

SKARLHVCQKI

Signal peptide:
amino acids 1-25

FIGURE 43

TCTCAATCTGCTGACCTCGTGATCCGCCTGACCTTGTAATCCACCTACCTTGGCCTCCCAAA

GTGTTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCAACATCACGTTTTTAAAAATTGATT

TCTTCAAATTCATGGCAAATATTTCCCTTCCCTTTAACTTCTTATGTCAGAATGAGGAAGGA

TAGCTGCATTTATTTAGTCAGTTTTCATTGCATAGTAATATTTTCATGTAGTATTTTCTAAG

TTATATTTTAGTAATTCATATGTTTTAGATTATAGGTTTTAACATACTTGTGAAAATACTTG

<u>ATG</u>TGTTTTAAAGCCTTGGGCAGAAATTCTGTATTGTTGAGGATTTGTTCTTTTATCCCCCT

TTTAAAGTCATCCGTCCTTGGCTCAGGATTTGGAGAGCTTGCACCACCAAAAATGGCAAACA

TCACCAGCTCCCAGATTTTGGACCAGTTGAAAGCTCCGAGTTTGGGCCAGTTTACCACCACC

CCAAGTACACAGCAGAATAGTACAAGTCACCCTACAACTACTACTTCTTGGGACCTCAAGCC

CCCAACATCCCAGTCCTCAGTCCTCAGTCATCTTGACTTCAAATCTCAACCTGAGCCATCCC

CAGTTCTTAGCCAGTTGAGCCAGCGACAACAGCACCAGAGCCAGGCAGTCACTGTTCCTCCT

CCTGGTTTGGAGTCCTTTCCTTCCCAGGCAAAACTTCGAGAATCAACACCTGGAGACAGTCC

CTCCACTGTGAACAAGCTTTTGCAGCTTCCCAGCACGACCATTGAAAATATCTCTGTGTCTG

TCCACCAGCCACAGCCCAAACACATCAAACTTGCTAAGCGGCGGATACCCCCAGCTTCTAAG

ATCCCAGCTTCTGCAGTGGAAATGCCTGGTTCAGCAGATGTCACAGGATTAAATGTGCAGTT

TGGGGCTCTGGAATTTGGGTCAGAACCTTCTCTCTCTGAATTTGGATCAGCTCCAAGCAGTG

AAAATAGTAATCAGATTCCCATCAGCTTGTATTCGAAGTCTTTAAGTGAGCCTTTGAATACA

TCTTTATCAATGACCAGTGCAGTACAGAACTCCACATATACAACTTCCGTCATTACCTCCTG

CAGTCTGACAAGCTCATCACTGAATTCTGCTAGTCCAGTAGCAATGTCTTCCTCTTATGACC

AGAGTTCTGTGCATAACAGGATCCCATACCAAAGCCCTGTGAGTTCATCAGAGTCAGCTCCA

GGAACCATCATGAATGGACATGGTGGTGGTCGAAGTCAGCAGACACTAGACAGTAAGTATAG

CAGCAAGCTACTCTTGTCATGGCTGGTGCCAACCAAACAGAGGAAGAGGATAGCTCACGTGA

TGTGGAAAACACCAGTTGGTCAATGGCTCATTCGT<u>TAA</u>AAAGCAGCCCTTTTGCTTTTTTGT

TTTTGGACCAGGTGTTGGCTGTGGTGTTATTAGAAATGTCTTAACCACAGCAAGAAGGAGGT

GGTGGTCTCATATTCTTCTGCCCTAATCAGACTGCACCACAAGTGCAGCATACAGTATGCAT

TTTAAAGATGCTTGGGCCAGGCGGGGTGGCTGATGCCCATAATCCCAGTGCTTTGGGGGGCC

AAGGCAGGCAGATTGCCCAAGCTCAGGAGTTTGAGACCACCCTGGGCAACATGGTGAAACTC

TGTCTCTACTAAAATACGAAAAACTAGCCGGGTGTGGTGGCGGCGCGTGCCTGTAATCCCAG

CTACTTGGGAGGCTGAGGCACAAGAATCGCTTGAGCCAGCTTGGGCTACAAAGTGAGACTCC

GTCTGAAAAGA

43/270

FIGURE 44

MCFKALGRNSVLLRICSFIPLLKSSVLGSGFGELAPPKMANITSSQILDQLKAPSLGQFTTT

PSTQQNSTSHPTTTTSWDLKPPTSQSSVLSHLDFKSQPEPSPVLSQLSQRQQHQSQAVTVPP

PGLESFPSQAKLRESTPGDSPSTVNKLLQLPSTTIENISVSVHQPQPKHIKLAKRRIPPASK

IPASAVEMPGSADVTGLNVQFGALEFGSEPSLSEFGSAPSSENSNQIPISLYSKSLSEPLNT

SLSMTSAVQNSTYTTSVITSCSLTSSSLNSASPVAMSSSYDQSSVHNRIPYQSPVSSSESAP

GTIMNGHGGGRSQQTLDSKYSSKLLLSWLVPTKQRKRIAHVMWKTPVGQWLIR

Signal peptide:
amino acids 1-24

FIGURE 45

GCCGAGTGGGACAAAGCCTGGGGCTGGGCGGGGGCC<u>ATG</u>GCGCTGCCATCCCGAATCCTGCT
TTGGAAACTTGTGCTTCTGCAGAGCTCTGCTGTTCTCCTGCACTCAGCGGTGGAGGAGACGG
ACGCGGGGCTGTACACCTGCAACCTGCACCATCACTACTGCCACCTCTACGAGAGCCTGGCC
GTCCGCCTGGAGGTCACCGACGGCCCCCCGGCCACCCCCGCCTACTGGGACGGCGAGAAGGA
GGTGCTGGCGGTGGCGCGCGGCGCACCCGCGCTTCTGACCTGCGTGAACCGCGGGCACGTGT
GGACCGACCGGCACGTGGAGGAGGCTCAACAGGTGGTGCACTGGGACCGGCAGCCGCCCGGG
GTCCCGCACGACCGCGCGGACCGCCTGCTGGACCTCTACGCGTCGGGCGAGCGCCGCGCCTA
CGGGCCCCTTTTTCTGCGCGACCGCGTGGCTGTGGGCGCGGATGCCTTTGAGCGCGGTGACT
TCTCACTGCGTATCGAGCCGCTGGAGGTCGCCGACGAGGGCACCTACTCCTGCCACCTGCAC
CACCATTACTGTGGCCTGCACGAACGCCGCGTCTTCCACCTGACGGTCGCCGAACCCCACGC
GGAGCCGCCCCCCGGGGCTCTCCGGGCAACGGCTCCAGCCACAGCGGCGCCCCAGGCCCAG
ACCCCACACTGGCGCGCGGCCACAACGTCATCAATGTCATCGTCCCCGAGAGCCGAGCCCAC
TTCTTCCAGCAGCTGGGCTACGTGCTGGCCACGCTGCTGCTCTTCATCCTGCTACTGGTCAC
TGTCCTCCTGGCCGCCCGCAGGCGCCGCGGAGGCTACGAATACTCGGACCAGAAGTCGGGAA
AGTCAAAGGGGGAAGGATGTTAACTTGGCGGAGTTCGCTGTGGCTGCAGGGGACCAGATGCTT
TACAGGAGTGAGGACATCCAGCTAGATTACAAAAACAACATCCTGAAGGAGAGGGCGGAGCT
GGCCCACAGCCCCCTGCCTGCCAAGTACATCGACCTAGACAAAGGGTTCCGGAAGGAGAACT
GCAAA<u>TAG</u>GGAGGCCCTGGGCTCCTGGCTGGGCCAGCAGCTGCACCTCTCCTGTCTGTGCTC
CTCGGGGCATCTCCTGATGCTCCGGGGCTCACCCCCCTTCCAGCGGCTGGTCCCGCTTTCCT
GGAATTTGGCCTGGGCGTATGCAGAGGCCGCCTCCACACCCCTCCCCCAGGGGCTTGGTGGC
AGCATAGCCCCCACCCCTGCGGCCTTTGCTCACGGGTGGCCCTGCCCACCCCTGGCACAACC
AAAATCCCACTGATGCCCATCATGCCCTCAGACCCTTCTGGGCTCTGCCCGCTGGGGGCCTG
AAGACATTCCTGGAGGACACTCCCATCAGAACCTGGCAGCCCCAAAACTGGGGTCAGCCTCA
GGGCAGGAGTCCCACTCCTCCAGGGCTCTGCTCGTCCGGGGCTGGGAGATGTTCCTGGAGGA
GGACACTCCCATCAGAACTTGGCAGCCTTGAAGTTGGGGTCAGCCTCGGCAGGAGTCCCACT
CCTCCTGGGGTGCTGCCTGCCACCAAGAGCTCCCCCACCTGTACCACCATGTGGGACTCCAG
GCACCATCTGTTCTCCCCAGGGACCTGCTGACTTGAATGCCAGCCCTTGCTCCTCTGTGTTG
CTTTGGGCCACCTGGGGCTGCACCCCCTGCCCTTTCTCTGCCCCATCCCTACCCTAGCCTTG
CTCTCAGCCACCTTGATAGTCACTGGGCTCCCTGTGACTTCTGACCCTGACACCCCTCCCTT
GGACTCTGCCTGGGCTGGAGTCTAGGGCTGGGGCTACATTTGGCTTCTGTACTGGCTGAGGA
CAGGGGAGGGAGTGAAGTTGGTTTGGGGTGGCCTGTGTTGCCACTCTCAGCACCCCACATTT
GCATCTGCTGGTGGACCTGCCACCATCACAATAAAGTCCCCATCTGATTTTTAAAAAAAAAA
AAAAA

FIGURE 46

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60618
<subunit 1 of 1, 341 aa, 1 stop
<MW: 38070, pI: 6.88, NX(S/T): 1
MALPSRILLWKLVLLQSSAVLLHSAVEETDAGLYTCNLHHHYCHLYESLAVRLEVTDGPPAT
PAYWDGEKEVLAVARGAPALLTCVNRGHVWTDRHVEEAQQVVHWDRQPPGVPHDRADRLLDL
YASGERRAYGPLFLRDRVAVGADAFERGDFSLRIEPLEVADEGTYSCHLHHHYCGLHERRVF
HLTVAEPHAEPPPRGSPGNGSSHSGAPGPDPTLARGHNVINVIVPESRAHFFQQLGYVLATL
LLFILLLVTVLLAARRRRGGYEYSDQKSGKSKGKDVNLAEFAVAAGDQMLYRSEDIQLDYKN
NILKERAELAHSPLPAKYIDLDKGFRKENCK
```

Important features:
Signal peptide:
amino acids 1-19

Transmembrane domain:
amino acids 237-262

N-glycosylation site.
amino acids 205-208

Cell attachment sequence.
amino acids 151-154

Coproporphyrinogen III oxidase proteins.
amino acids 115-140

46/270

FIGURE 47

```
CGCCGGAGGCAGCGGCGGCGTGGCGCAGCGGCGACATGGCCGTTGTCTCAGAGGACGACTTT
CAGCACAGTTCAAACTCCACCTACGGAACCACAAGCAGCAGTCTCCGAGCTGACCAGGAGGC
ACTGCTTGAGAAGCTGCTGGACCGCCCGCCCCCTGGCCTGCAGAGGCCCGAGGACCGCTTCT
GTGGCACATACATCATCTTCTTCAGCCTGGGCATTGGCAGTCTACTGCCATGGAACTTCTTT
ATCACTGCCAAGGAGTACTGGATGTTCAAACTCCGCAACTCCTCCAGCCCAGCCACCGGGGA
GGACCCTGAGGGCTCAGACATCCTGAACTACTTTGAGAGCTACCTTGCCGTTGCCTCCACCG
TGCCCTCCATGCTGTGCCTGGTGGCCAACTTCCTGCTTGTCAACAGGGTTGCAGTCCACATC
CGTGTCCTGGCCTCACTGACGGTCATCCTGGCCATCTTCATGGTGATAACTGCACTGGTGAA
GGTGGACACTTCCTCCTGGACCCGTGGTTTTTTTGCGGTCACCATTGTCTGCATGGTGATCC
TCAGCGGTGCCTCCACTGTCTTCAGCAGCAGCATCTACGGCATGACCGGCTCCTTTCCTATG
AGGAACTCCCAAGCACTGATATCAGGAGGAGCCATGGGCGGGACGGTCAGCGCCGTGGCCTC
ATTGGTGGACTTGGCTGCATCCAGTGATGTGAGGAACAGCGCCCTGGCCTTCTTCCTGACGG
CCACCATCTTCCTCGTGCTCTGCATGGGACTCTACCTGCTGCTGTCCAGGCTGGAGTATGCC
AGGTACTACATGAGGCCTGTTCTTGCGGCCCATGTGTTTTCTGGTGAAGAGGAGCTTCCCCA
GGACTCCCTCAGTGCCCCTTCGGTGGCCTCCAGATTCATTGATTCCCACACACCCCCTCTCC
GCCCCATCCTGAAGAAGACGGCCAGCCTGGGCTTCTGTGTCACCTACGTCTTCTTCATCACC
AGCCTCATCTACCCCGCCGTCTGCACCAACATCGAGTCCCTCAACAAGGGCTCGGGCTCACT
GTGGACCACCAAGTTTTTCATCCCCCTCACTACCTTCCTCCTGTACAACTTTGCTGACCTAT
GTGGCCGGCAGCTCACCGCCTGGATCCAGGTGCCAGGGCCCAACAGCAAGGCGCTCCCAGGG
TTCGTGCTCCTCCGGACCTGCCTCATCCCCCTCTTCGTGCTCTGTAACTACCAGCCCCGCGT
CCACCTGAAGACTGTGGTCTTCCAGTCCGATGTGTACCCCGCACTCCTCAGCTCCCTGCTGG
GGCTCAGCAACGGCTACCTCAGCACCCTGGCCCTCCTCTACGGGCCTAAGATTGTGCCCAGG
GAGCTGGCTGAGGCCACGGGAGTGGTGATGTCCTTTTATGTGTGCTTGGGCTTAACACTGGG
CTCAGCCTGCTCTACCCTCCTGGTGCACCTCATCTAGAAGGGAGGACACAAGGACATTGGTG
CTTCAGAGCCTTTGAAGATGAGAAGAGAGTGCAGGAGGGCTGGGGGCCATGGAGGAAAGGCC
TAAAGTTTCACTTGGGGACAGAGAGCAGAGCACACTCGGGCCTCATCCCTCCCAAGATGCCA
GTGAGCCACGTCCATGCCCATTCCGTGCAAGGCAGATATTCCAGTCATATTAACAGAACACT
CCTGAGACAGTTGAAGAAGAAATAGCACAAATCAGGGGTACTCCCTTCACAGCTGATGGTTA
ACATTCCACCTTCTTTCTAGCCCTTCAAAGATGCTGCCAGTGTTCGCCCTAGAGTTATTACA
AAGCCAGTGCCAAAACCCAGCCATGGGCTCTTTGCAACCTCCCAGCTGCGCTCATTCCAGCT
GACAGCGAGATGCAAGCAAATGCTCAGCTCTCCTTACCCTGAAGGGGTCTCCCTGGAATGGA
AGTCCCCTGGCATGGTCAGTCCTCAGGCCCAAGACTCAAGTGTGCACAGACCCCTGTGTTCT
GCGGGTGAACAACTGCCCACTAACCAGACTGGAAAACCCAGAAAGATGGGCCTTCCATGAAT
GCTTCATTCCAGAGGGACCAGAGGGCCTCCCTGTGCAAGGGATCAAGCATGTCTGGCCTGGG
TTTTCAAAAAAAGAGGGATCCTCATGACCTGGTGGTCTATGGCCTGGGTCAAGATGAGGGTC
TTTCAGTGTTCCTGTTTACAACATGTCAAAGCCATTGGTTCAAGGGCGTAATAAATACTTGC
GTATTCAAAAA
```

FIGURE 48

MAVVSEDDFQHSSNSTYGTTSSSLRADQEALLEKLLDRPPPGLQRPEDRFCGTYIIFFSLGI

GSLLPWNFFITAKEYWMFKLRNSSSPATGEDPEGSDILNYFESYLAVASTVPSMLCLVANFL

LVNRVAVHIRVLASLTVILAIFMVITALVKVDTSSWTRGFFAVTIVCMVILSGASTVFSSSI

YGMTGSFPMRNSQALISGGAMGGTVSAVASLVDLAASSDVRNSALAFFLTATIFLVLCMGLY

LLLSRLEYARYYMRPVLAAHVFSGEEELPQDSLSAPSVASRFIDSHTPPLRPILKKTASLGF

CVTYVFFITSLIYPAVCTNIESLNKGSGSLWTTKFFIPLTTFLLYNFADLCGRQLTAWIQVP

GPNSKALPGFVLLRTCLIPLFVLCNYQPRVHLKTVVFQSDVYPALLSSLLGLSNGYLSTLAL

LYGPKIVPRELAEATGVVMSFYVCLGLTLGSACSTLLVHLI

Transmembrane domain:
amino acids 50-74 (type II), 105-127, 135-153, 163-183, 228-252,
305-330, 448-472

FIGURE 49

GACAGTGGAGGGCAGTGGAGAGGACCGCGCTGTCCTGCTGTCACCAAGAGCTGGAGACACCA

TCTCCCACCGAGAGTC<u>ATG</u>GCCCCATTGGCCCTGCACCTCCTCGTCCTCGTCCCCATCCTCC

TCAGCCTGGTGGCCTCCCAGGACTGGAAGGCTGAACGCAGCCAAGACCCCTTCGAGAAATGC

ATGCAGGATCCTGACTATGAGCAGCTGCTCAAGGTGGTGACCTGGGGGCTCAATCGGACCCT

GAAGCCCCAGAGGGTGATTGTGGTTGGCGCTGGTGTGGCCGGGCTGGTGGCCGCCAAGGTGC

TCAGCGATGCTGGACACAAGGTCACCATCCTGGAGGCAGATAACAGGATCGGGGGCCGCATC

TTCACCTACCGGGACCAGAACACGGGCTGGATTGGGGAGCTGGGAGCCATGCGCATGCCCAG

CTCTCACAGGATCCTCCACAAGCTCTGCCAGGGCCTGGGGCTCAACCTGACCAAGTTCACCC

AGTACGACAAGAACACGTGGACGGAGGTGCACGAAGTGAAGCTGCGCAACTATGTGGTGGAG

AAGGTGCCCGAGAAGCTGGGCTACGCCTTGCGTCCCCAGGAAAAGGGCCACTCGCCCGAAGA

CATCTACCAGATGGCTCTCAACCAGGCCCTCAAAGACCTCAAGGCACTGGGCTGCAGAAAGG

CGATGAAGAAGTTTGAAAGGCACACGCTCTTGGAATATCTTCTCGGGGAGGGGAACCTGAGC

CGGCCGGCCGTGCAGCTTCTGGGAGACGTGATGTCCGAGGATGGCTTCTTCTATCTCAGCTT

CGCCGAGGCCCTCCGGGCCCACAGCTGCCTCAGCGACAGACTCCAGTACAGCCGCATCGTGG

GTGGCTGGGACCTGCTGCCGCGCGCGCTGCTGAGCTCGCTGTCCGGGCTTGTGCTGTTGAAC

GCGCCCGTGGTGGCGATGACCCAGGGACCGCACGATGTGCACGTGCAGATCGAGACCTCTCC

CCCGGCGCGGAATCTGAAGGTGCTGAAGGCCGACGTGGTGCTGCTGACGGCGAGCGGACCGG

CGGTGAAGCGCATCACCTTCTCGCCGCCGCTGCCCCGCCACATGCAGGAGGCGCTGCGGAGG

CTGCACTACGTGCCGGCCACCAAGGTGTTCCTAAGCTTCCGCAGGCCCTTCTGGCGCGAGGA

GCACATTGAAGGCGGCCACTCAAACACCGATCGCCCGTCGCGCATGATTTTCTACCGCCGC

CGCGCGAGGGCGCGCTGCTGCTGGCCTCGTACACGTGGTCGGACGCGGCGGCAGCGTTCGCC

GGCTTGAGCCGGGAAGAGGCGTTGCGCTTGGCGCTCGACGACGTGGCGGCATTGCACGGGCC

TGTCGTGCGCCAGCTCTGGGACGGCACCGGCGTCGTCAAGCGTTGGGCGGAGGACCAGCACA

GCCAGGGTGGCTTTGTGGTACAGCCGCCGGCGCTCTGGCAAACCGAAAAGGATGACTGGACG

GTCCCTTATGGCCGCATCTACTTTGCCGGCGAGCACACCGCCTACCCGCACGGCTGGGTGGA

GACGGCGGTCAAGTCGGCGCTGCGCGCCGCCATCAAGATCAACAGCCGGAAGGGGCCTGCAT

CGGACACGGCCAGCCCCGAGGGGCACGCATCTGACATGGAGGGGCAGGGGCATGTGCATGGG

GTGGCCAGCAGCCCCTCGCATGACCTGGCAAAGGAAGAAGGCAGCCACCCTCCAGTCCAAGG

CCAGTTATCTCTCCAAAACACGACCCACACGAGGACCTCGCAT<u>TAA</u>AGTATTTTCGGAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

49/270

FIGURE 50

MAPLALHLLVLVPILLSLVASQDWKAERSQDPFEKCMQDPDYEQLLKVVTWGLNRTLKPQRV

IVVGAGVAGLVAAKVLSDAGHKVTILEADNRIGGRIFTYRDQNTGWIGELGAMRMPSSHRIL

HKLCQGLGLNLTKFTQYDKNTWTEVHEVKLRNYVVEKVPEKLGYALRPQEKGHSPEDIYQMA

LNQALKDLKALGCRKAMKKFERHTLLEYLLGEGNLSRPAVQLLGDVMSEDGFFYLSFAEALR

AHSCLSDRLQYSRIVGGWDLLPRALLSSLSGLVLLNAPVVAMTQGPHDVHVQIETSPPARNL

KVLKADVVLLTASGPAVKRITFSPPLPRHMQEALRRLHYVPATKVFLSFRRPFWREEHIEGG

HSNTDRPSRMIFYPPPREGALLLASYTWSDAAAAFAGLSREEALRLALDDVAALHGPVVRQL

WDGTGVVKRWAEDQHSQGGFVVQPPALWQTEKDDWTVPYGRIYFAGEHTAYPHGWVETAVKS

ALRAAIKINSRKGPASDTASPEGHASDMEGQGHVHGVASSPSHDLAKEEGSHPPVQGQLSLQ

NTTHTRTSH

Signal peptide:
amino acids 1-21

50/270

FIGURE 51

CTGACATGGCCTGACTCGGGACAGCTCAGAGCAGGGCAGAACTGGGGACACTCTGGGCCGGCCTTCTG
CCTGCATGGACGCTCTGAAGCCACCCTGTCTCTGGAGGAACCACGAGCGAGGGAAGAAGGACAGGGAC
TCGTGTGGCAGGAAGAACTCAGAGCCGGGAAGCCCCCATTCACTAGAAGCACTGAGAGATGCGGCCCC
CTCGCAGGGTCTGAATTTCCTGCTGCTGTTCACAAAGATGCTTTTTATCTTTAACTTTTTGTTTTCCC
CACTTCCGACCCCGGCGTTGATCTGCATCCTGACATTTGGAGCTGCCATCTTCTTGTGGCTGATCACC
AGACCTCAACCCGTCTTACCTCTTCTTGACCTGAACAATCAGTCTGTGGGAATTGAGGGAGGAGCACG
GAAGGGGGTTTCCCAGAAGAACAATGACCTAACAAGTTGCTGCTTCTCAGATGCCAAGACTATGTATG
AGGTTTTCCAAAGAGGACTCGCTGTGTCTGACAATGGGCCCTGCTTGGGATATAGAAAACCAAACCAG
CCCTACAGATGGCTATCTTACAAACAGGTGTCTGATAGAGCAGAGTACCTGGGTTCCTGTCTCTTGCA
TAAAGGTTATAAATCATCACCAGACCAGTTTGTCGGCATCTTTGCTCAGAATAGGCCAGAGTGGATCA
TCTCCGAATTGGCTTGTTACACGTACTCTATGGTAGCTGTACCTCTGTATGACACCTTGGGACCAGAA
GCCATCGTACATATTGTCAACAAGGCTGATATCGCCATGGTGATCTGTGACACACCCCAAAAGGCATT
GGTGCTGATAGGGAATGTAGAGAAAGGCTTCACCCCGAGCCTGAAGGTGATCATCCTTATGGACCCCT
TTGATGATGACCTGAAGCAAAGAGGGGAGAAGAGTGGAATTGAGATCTTATCCCTATATGATGCTGAG
AACCTAGGCAAAGAGCACTTCAGAAAACCTGTGCCTCCTAGCCCAGAAGACCTGAGCGTCATCTGCTT
CACCAGTGGGACCACAGGTGACCCCAAAGGAGCCATGATAACCCATCAAAATATTGTTTCAAATGCTG
CTGCCTTTCTCAAATGTGTGGAGCATGCTTATGAGCCCACTCCTGATGATGTGGCCATATCCTACCTC
CCTCTGGCTCATATGTTTGAGAGGATTGTACAGGCTGTTGTGTACAGCTGTGGAGCCAGAGTTGGATT
CTTCCAAGGGGATATTCGGTTGCTGGCTGACGACATGAAGACTTTGAAGCCCACATTGTTTCCCGCGG
TGCCTCGACTCCTTAACAGGATCTACGATAAGGTACAAAATGAGGCCAAGACACCCTTGAAGAAGTTC
TTGTTGAAGCTGGCTGTTTCCAGTAAATTCAAAGAGCTTCAAAAGGGTATCATCAGGCATGATAGTTT
CTGGGACAAGCTCATCTTTGCAAAGATCCAGGACAGCCTGGGCGGAAGGGTTCGTGTAATTGTCACTG
GAGCTGCCCCCATGTCCACTTCAGTCATGACATTCTTCCGGGCAGCAATGGGATGTCAGGTGTATGAA
GCTTATGGTCAAACAGAATGCACAGGTGGCTGTACATTTACATTACCTGGGGACTGGACATCAGGTCA
CGTTGGGGTGCCCCTGGCTTGCAATTACGTGAAGCTGGAAGATGTGGCTGACATGAACTACTTTACAG
TGAATAATGAGGAGAGGTCTGCATCAAGGGTACAAACGTGTTCAAAGGATACCTGAAGGACCCTGAG
AAGACACAGGAAGCCCTGGACAGTGATGGCTGGCTTCACACAGGAGACATTGGTCGCTGGCTCCCGAA
TGGAACTCTGAAGATCATCGACCGTAAAAAGAACATTTTCAAGCTGGCCCAAGGAGAATACATTGCAC
CAGAGAAGATAGAAAATATCTACAACAGGAGTCAACCAGTGTTACAAATTTTTGTACACGGGGAGAGC
TTACGGTCATCCTTAGTAGGAGTGGTGGTTCCTGACACAGATGTACTTCCCTCATTTGCAGCCAAGCT
TGGGGTGAAGGGCTCCTTTGAGGAACTGTGCCAAAACCAAGTTGTAAGGGAAGCCATTTTAGAAGACT
TGCAGAAAATTGGGAAAGAAAGTGGCCTTAAAACTTTTGAACAGGTCAAAGCCATTTTTTCTTCATCCA
GAGCCATTTTCCATTGAAAATGGGCTCTTGACACCAACATTGAAAGCAAAGCGAGGAGAGCTTTCCAA
ATACTTTCGGACCCAAATTGACAGCCTGTATGAGCACATCCAGGATTAGGATAAGGTACTTAAGTACC
TGCCGGCCCACTGTGCACTGCTTGTGAGAAAATGGATTAAAAACTATTCTTACATTTGTTTTGCCTTT
CCTCCTATTTTTTTTTTAACCTGTTAAACTCTAAAGCCATAGCTTTTGTTTTATATTGAGACATATAAT
GTGTAAACTTAGTTCCCAAATAAATCAATCCTGTCTTTCCCATCTTCGATGTTGCTAATATTAAGGCT
TCAGGGCTACTTTTATCAACATGCCTGTCTTCAAGATCCCAGTTTATGTTCTGTGTCCTTCCTCATGA
TTTCCAACCTTAATACTATTAGTAACCACAAGTTCAAGGGTCAAAGGGACCCTCTGTGCCTTCTTCTT
TGTTTTGTGATAAACATAACTTGCCAACAGTCTCTATGCTTATTTACATCTTCTACTGTTCAAACTAA
GAGATTTTTAAATTCTGAAAAACTGCTTACAATTCATGTTTTCTAGCCACTCCACAAACCACTAAAAT
TTTAGTTTTAGCCTATCACTCATGTCAATCATATCTATGAGACAAATGTCTCCGATGCTCTTCTGCGT
AAATTAAATTGTGTACTGAAGGGAAAAGTTTGATCATACCAAACATTTCCTAAACTCTCTCAGTTAGAT
ATCTGACTTGGGAGTATTAAAAATTGGGTCTATGACATACTGTCCAAAAGGAATGCTGTTCTTAAAGC
ATTATTTACAGTAGGAACTGGGGAGTAAATCTGTTCCCTACAGTTTGCTGCTGAGCTGGAAGCTGTGG
GGGAAGGAGTTGACAGGTGGGCCCAGTGAACTTTTCCAGTAAATGAAGCAAGCACTGAATAAAAACCT
CCTGAACTGGGAACAAAGATCTACAGGCAAGCAAGATGCCCACACAACAGGCTTATTTTCTGTGAAGG
AACCAACTGATCTCCCCCACCCTTGGATTAGAGTTCCTGCTCTACCTTACCCACAGATAACACATGTT
GTTTCTACTTGTAAATGTAAAGTCTTTAAAATAAACTATTACAGATAAAAAA

FIGURE 52

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60775
<subunit 1 of 1, 739 aa, 1 stop
<MW: 82263, pI: 7.55, NX(S/T): 3
MDALKPPCLWRNHERGKKDRDSCGRKNSEPGSPHSLEALRDAAPSQGLNFLLLFTKMLFIFN
FLFSPLPTPALICILTFGAAIFLWLITRPQPVLPLLDLNNQSVGIEGGARKGVSQKNNDLTS
CCFSDAKTMYEVFQRGLAVSDNGPCLGYRKPNQPYRWLSYKQVSDRAEYLGSCLLHKGYKSS
PDQFVGIFAQNRPEWIISELACYTYSMVAVPLYDTLGPEAIVHIVNKADIAMVICDTPQKAL
VLIGNVEKGFTPSLKVIILMDPFDDDLKQRGEKSGIEILSLYDAENLGKEHFRKPVPPSPED
LSVICFTSGTTGDPKGAMITHQNIVSNAAAFLKCVEHAYEPTPDDVAISYLPLAHMFERIVQ
AVVYSCGARVGFFQGDIRLLADDMKTLKPTLFPAVPRLLNRIYDKVQNEAKTPLKKFLLKLA
VSSKFKELQKGIIRHDSFWDKLIFAKIQDSLGGRVRVIVTGAAPMSTSVMTFFRAAMGCQVY
EAYGQTECTGGCTFTLPGDWTSGHVGVPLACNYVKLEDVADMNYFTVNNEGEVCIKGTNVFK
GYLKDPEKTQEALDSDGWLHTGDIGRWLPNGTLKIIDRKKNIFKLAQGEYIAPEKIENIYNR
SQPVLQIFVHGESLRSSLVGVVVPDTDVLPSFAAKLGVKGSFEELCQNQVVREAILEDLQKI
GKESGLKTFEQVKAIFLHPEPFSIENGLLTPTLKAKRGELSKYFRTQIDSLYEHIQD
```

Important features:

Type II transmembrane domain:
amino acids 61-80

Putative AMP-binding domain signature.
amino acids 314-325

N-glycosylation site.
amino acids 102-105, 588-591 and 619-622

FIGURE 53

```
GGAGGCGGAGGCCGCGGCGAGCCGGGCCGAGCAGTGAGGGCCCTAGCGGGGCCCGAGCGGGG
CCCGGGGCCCCTAAGCCATTCCTGAAGTCATGGGCTGGCCAGGACATTGGTGACCCGCCAAT
CCGGTATGGACGACTGGAAGCCCAGCCCCCTCATCAAGCCCTTTGGGGCTCGGAAGAAGCGG
AGCTGGTACCTTACCTGGAAGTATAAACTGACAAACCAGCGGGCCCTGCGGAGATTCTGTCA
GACAGGGGCCGTGCTTTTCCTGCTGGTGACTGTCATTGTCAATATCAAGTTGATCCTGGACA
CTCGGCGAGCCATCAGTGAAGCCAATGAAGACCCAGAGCCAGAGCAAGACTATGATGAGGCC
CTAGGCCGCCTGGGAGCCCCCACGGCGCAGAGGCAGTGGTCCCCGGCGGGTCCTGGACGTAGA
GGTGTATTCAAGTCGCAGCAAAGTATATGTGGCAGTGGATGGCACCACGGTGCTGGAGGATG
AGGCCCGGGAGCAGGGCCGGGGCATCCATGTCATTGTCCTCAACCAGGCCACGGGCCACGTG
ATGGCAAAACGTGTGTTTGACACGTACTCACCTCATGAGGATGAGGCCATGGTGCTATTCCT
CAACATGGTAGCGCCCGGCCGAGTGCTCATCTGCACTGTCAAGGATGAGGGCTCCTTCCACC
TCAAGGACACAGCCAAGGCTCTGCTGAGGAGCCTGGGCAGCCAGGCTGGCCCTGCCCTGGGC
TGGAGGGACACATGGGCCTTCGTGGGACGAAAAGGAGGTCCTGTCTTCGGGGAGAAACATTC
TAAGTCACCTGCCCTCTCTTCCTGGGGGGGACCCAGTCCTGCTGAAGACAGATGTGCCATTGA
GCTCAGCAGAAGAGGCAGAGTGCCACTGGGCAGACACAGAGCTGAACCGTCGCCGCCGGCGC
TTCTGCAGCAAAGTTGAGGGCTATGGAAGTGTATGCAGCTGCAAGGACCCCACACCCATCGA
GTTCAGCCCTGACCCACTCCCAGACAACAAGGTCCTCAATGTGCCTGTGGCTGTCATTGCAG
GGAACCGACCCAATTACCTGTACAGGATGCTGCGCTCTCTGCTTTCAGCCCAGGGGGTGTCT
CCTCAGATGATAACAGTTTTCATTGACGGCTACTATGAGGAACCCATGGATGTGGTGGCACT
GTTTGGTCTGAGGGGCATCCAGCATACTCCCATCAGCATCAAGAATGCCCGCGTGTCTCAGC
ACTACAAGGCCAGCCTCACTGCCACTTTCAACCTGTTTCCGGAGGCCAAGTTTGCTGTGGTT
CTGGAAGAGGACCTGGACATTGCTGTGGATTTTTTTCAGTTTCCTGAGCCAATCCATCCACCT
ACTGGAGGAGGATGACAGCCTGTACTGCATCTCTGCCTGGAATGACCAGGGGTATGAACACA
CGGCTGAGGACCCAGCACTACTGTACCGTGTGGAGACCATGCCTGGGCTGGGCTGGGTGCTC
AGGAGGTCCTTGTACAAGGAGGAGCTTGAGCCCAAGTGGCCTACACCGGAAAAGCTCTGGGA
TTGGGACATGTGGATGCGGATGCCTGAACAACGCCGGGGCCGAGAGTGCATCATCCCTGACG
TTTCCCGATCCTACCACTTTGGCATCGTCGGCCTCAACATGAATGGCTACTTTCACGAGGCC
TACTTCAAGAAGCACAAGTTCAACACGGTTCCAGGTGTCCAGCTCAGGAATGTGGACAGTCT
GAAGAAAGAAGCTTATGAAGTGGAAGTTCACAGGCTGCTCAGTGAGGCTGAGGTTCTGGACC
ACAGCAAGAACCCTTGTGAAGACTCTTTCCTGCCAGACACAGAGGGCCACACCTACGTGGCC
TTTATTCGAATGGAGAAAGATGATGACTTCACCACCTGGACCCAGCTTGCCAAGTGCCTCCA
TATCTGGGACCTGGATGTGCGTGGCAACCATCGGGGCCTGTGGAGATTGTTTCGGAAGAAGA
ACCACTTCCTGGTGGTGGGGGTCCCGGCTTCCCCCCTACTCAGTGAAGAAGCCACCCTCAGTC
ACCCCAATTTTCCTGGAGCCACCCCCAAAGGAGGAGGGAGCCCCAGGAGCCCCAGAACAGAC
ATGAGACCTCCTCCAGGACCCTGCGGGGCTGGGTACTGTGTACCCCCAGGCTGGCTAGCCCT
TCCCTCCATCCTGTAGGATTTTGTAGATGCTGGTAGGGGCTGGGGCTACCTTGTTTTTAACA
TGAGACTTAATTACTAACTCCAAGGGGAGGGTTCCCCTGCTCCAACACCCCGTTCCTGAGTT
AAAAGTCTATTTATTTACTTCCTTGTTGGAGAAGGGCAGGAGAGTACCTGGGAATCATTACG
ATCCCTAGCAGCTCATCCTGCCCTTTGAATACCCTCACTTTCCAGGCCTGGCTCAGAATCTA
ACCTATTTATTGACTGTCCTGAGGGCCTTGAAAACAGGCCGAACCTGGAGGGCCTGGATTTC
TTTTTGGGCTGGAATGCTGCCCTGAGGGTGGGGCTGGCTCTTACTCAGGAAACTGCTGTGCC
CAACCCATGGACAGGCCCAGCTGGGGCCCACATGCTGACACAGACTCACTCAGAGACCCTTA
GACACTGGACCAGGCCTCCTCTCAGCCTTCTCTTTGTCCAGATTTCCAAAGCTGGATAAGTT
GGTCATTGATTAAAAAAGGAGAAGCCCTCTGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

53/270

FIGURE 54

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA61185
><subunit 1 of 1, 660 aa, 1 stop
><MW: 75220, pI: 6.76, NX(S/T): 0
MDDWKPSPLIKPFGARKKRSWYLTWKYKLTNQRALRRFCQTGAVLFLLVTVIVNIKLILDTR
RAISEANEDPEPEQDYDEALGRLEPPRRRGSGPRRVLDVEVYSSRSKVYVAVDGTTVLEDEA
REQGRGIHVIVLNQATGHVMAKRVFDTYSPHEDEAMVLFLNMVAPGRVLICTVKDEGSFHLK
DTAKALLRSLGSQAGPALGWRDTWAFVGRKGGPVFGEKHSKSPALSSWGDPVLLKTDVPLSS
AEEAECHWADTELNRRRRRFCSKVEGYGSVCSCKDPTPIEFSPDPLPDNKVLNVPVAVIAGN
RPNYLYRMLRSLLSAQGVSPQMITVFIDGYYEEPMDVVALFGLRGIQHTPISIKNARVSQHY
KASLTATFNLFPEAKFAVVLEEDLDIAVDFFSFLSQSIHLLEEDDSLYCISAWNDQGYEHTA
EDPALLYRVETMPGLGWVLRRSLYKEELEPKWPTPEKLWDWDMWMRMPEQRRGRECIIPDVS
RSYHFGIVGLNMNGYFHEAYFKKHKFNTVPGVQLRNVDSLKKEAYEVEVHRLLSEAEVLDHS
KNPCEDSFLPDTEGHTYVAFIRMEKDDDFTTWTQLAKCLHIWDLDVRGNHRGLWRLFRKKNH
FLVVGVPASPYSVKKPPSVTPIFLEPPPKEEGAPGAPEQT
```

Important features of the protein:
Transmembrane domain:
amino acids 38-55

Homologous region to Mouse GNT1
amino acids 229-660

FIGURE 55

```
CGGACGCGTGGGCTGCTGGTGGGAAGGCCTAAAGAACTGGAAAGCCCACTCTCTTGGAACCACCACAC
CTGTTTAAAGAACCTAAGCACCATTTAAAGCCACTGGAAATTTGTTGTCTAGTGGTTGTGGGTGAATA
AAGGAGGGCAGAATGGATGATTTCATCTCCATTAGCCTGCTGTCTCTGGCTATGTTGGTGGGATGTTA
CGTGGCCGGAATCATTCCCTTGGCTGTTAATTTCTCAGAGGAACGACTGAAGCTGGTGACTGTTTTGG
GTGCTGGCCTTCTCTGTGGAACTGCTCTGGCAGTCATCGTGCCTGAAGGAGTACATGCCCTTTATGAA
GATATTCTTGAGGGAAAACACCACCAAGCAAGTGAAACACATAATGTGATTGCATCAGACAAAGCAGC
AGAAAAATCAGTTGTCCATGAACATGAGCACAGCCACGACCACACACAGCTGCATGCCTATATTGGTG
TTTCCCTCGTTCTGGGCTTCGTTTTCATGTTGCTGGTGGACCAGATTGGTAACTCCCATGTGCATTCT
ACTGACGATCCAGAAGCAGCAAGGTCTAGCAATTCCAAAATCACCACCACGCTGGGTCTGGTTGTCCA
TGCTGCAGCTGATGGTGTTGCTTTGGGAGCAGCAGCATCTACTTCACAGACCAGTGTCCAGTTAATTG
TGTTTGTGGCAATCATGCTACATAAGGCACCAGCTGCTTTTGGACTGGTTTCCTTCTTGATGCATGCT
GGCTTAGAGCGGAATCGAATCAGAAAGCACTTGCTGGTCTTTGCATTGGCAGCACCAGTTATGTCCAT
GGTGACATACTTAGGACTGAGTAAGAGCAGTAAAGAAGCCCTTTCAGAGGTGAACGCCACGGGAGTGG
CCATGCTTTTCTCTGCCGGGACATTTCTTTATGTTGCCACAGTACATGTCCTCCCTGAGGTGGGCGGA
ATAGGGCACAGCCACAAGCCCGATGCCACGGGAGGGAGAGGCCTCAGCCGCCTGGAAGTGGCAGCCCT
GGTTCTGGGTTGCCTCATCCCTCTCATCCTGTCAGTAGGACACCAGCATTAAATGTTCAAGGTCCAGC
CTTGGTCCAGGGCCGTTTGCCATCCAGTGAGAACAGCCGGCACGTGACAGCTACTCACTTCCTCAGTC
TCTTGTCTCACCTTGCGCATCTCTACATGTATTCCTAGAGTCCAGAGGGGAGGTGAGGTTAAAACCTG
AGTAATGGAAAAGCTTTTAGAGTAGAAACACATTTACGTTGCAGTTAGCTATAGACATCCCATTGTGT
TATCTTTTAAAAGGCCCTTGACATTTTGCGTTTTAATATTTCTCTTAACCCTATTCTCAGGGAAGATG
GAATTTAGTTTTAAGGAAAAGAGGAGAACTTCATACTCACAATGAAATAGTGATTATGAAAATACAGT
GTTCTGTAATTAAGCTATGTCTCTTTCTTCTTAGTTTAGAGGCTCTGCTACTTTATCCATTGATTTTT
AACATGGTTCCCACCATGTAAGACTGGTGCTTTAGCATCTATGCCACATGCGTTGATGGAAGGTCATA
GCACCCACTCACTTAGATGCTAAAGGTGATTCTAGTTAATCTGGGATTAGGGTCAGGAAAATGATAGC
AAGACACATTGAAAGCTCTCTTTATACTCAAAAGAGATATCCATTGAAAAGGGATGTCTAGAGGGATT
TAAACAGCTCCTTTGGCACGTGCCTCTCTGAATCCAGCCTGCCATTCCATCAAATGGAGCAGGAGAGG
TGGGAGGAGCTTCTAAAGAGGTGACTGGTATTTTGTAGCATTCCTTGTCAAGTTCTCCTTTGCAGAAT
ACCTGTCTCCACATTCCTAGAGAGGAGCCAAGTTCTAGTAGTTTCAGTTCTAGGCTTTCCTTCAAGAA
CAGTCAGATCACAAAGTGTCTTTGGAAATTAAGGGATATTAAATTTTAAGTGATTTTTGGATGGTTAT
TGATATCTTTGTAGTAGCTTTTTTTAAAAGACTACCAAAATGTATGGTTGTCCTTTTTTTTTTGTTTTT
TTTTTTTTTAATTATTTCTCTTAGCAGATCAGCAATCCCTCTAGGGACCTAAATACTAGGTCAGCTTT
GGCGACACTGTGTCTTCTCACATAACCACCTGTAGCAAGATGGATCATAAATGAGAAGTGTTTGCCTA
TTGATTTAAAGCTTATTGGAATCATGTCTCTTGTCTCTTCGTCTTTTCTTTGCTTTTCTTCTAACTTT
TCCCTCTAGCCTCTCCTCGCCACAATTTGCTGCTTACTGCTGGTGTTAATATTTGTGTGGGATGAATT
CTTATCAGGACAACCACTTCTCGAACTGTAATAATGAAGATAATAATATCTTTATTCTTTATCCCCTT
CAAAGAAATTACCTTTGTGTCAAATGCCGCTTTGTTGAGCCCTTAAAATACCACCTCCTCATGTGTAA
ATTGACACAATCACTAATCTGGTAATTTAAACAATTGAGATAGCAAAAGTGTTTAACAGACTAGGATA
ATTTTTTTTTCATATTTGCCAAAATTTTTGTAAACCCTGTCTTGTCAAATAAGTGTATAATATTGTAT
TATTAATTTATTTTTACTTTCTATACCATTTCAAAACACATTACACTAAGGGGGAACCAAGACTAGTT
TCTTCAGGGCAGTGGACGTAGTAGTTTGTAAAAACGTTTTCTATGACGCATAAGCTAGCATGCCTATG
ATTTATTTCCTTCATGAATTTGTCACTGGATCAGCAGCTGTGGAAATAAAGCTTGTGAGCCCTCTGCT
GGCCACAGTGAGGAAAGTAGCACAAATAGGATACAGTTGTATGTAGTCATTGGCAACAATTGCATACA
ATTTTACTACCAAGAGAAGGTATAGTATGGAAAGTCCAAATGACTTCCTTGATTGGATGTTAACAGCT
GACTGGTGTGAGACTTGAGGTTTCATCTAGTCCTTCAAAACTATATGGTTGCCTAGATTCTCTCTGGA
AACTGACTTTGTCAAATAAATAGCAGATTGTAGTGTCAAAAAAAA
```

55 1270

FIGURE 56

MDDFISISLLSLAMLVGCYVAGIIPLAVNFSEERLKLVTVLGAGLLCGTALAVIVPEGVHAL

YEDILEGKHHQASETHNVIASDKAAEKSVVHEHEHSHDHTQLHAYIGVSLVLGFVFMLLVDQ

IGNSHVHSTDDPEAARSSNSKITTTLGLVVHAAADGVALGAAASTSQTSVQLIVFVAIMLHK

APAAFGLVSFLMHAGLERNRIRKHLLVFALAAPVMSMVTYLGLSKSSKEALSEVNATGVAML

FSAGTFLYVATVHVLPEVGGIGHSHKPDATGGRGLSRLEVAALVLGCLIPLILSVGHQH

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 37-56, 106-122, 211-230, 240-260, 288-304

FIGURE 57

```
GCTCGAGGCCGGCGGCGGCGGGAGAGCGACCCGGGCGGCCTCGTAGCGGGGCCCCGGATCCC
CGAGTGGCGGCCGGAGCCTCGAAAAGAGATTCTCAGCGCTGATTTTGAGATGATGGGCTTGG
GAAACGGGCGTCGCAGCATGAAGTCGCCGCCCCTCGTGCTGGCCGCCCTGGTGGCCTGCATC
ATCGTCTTGGGCTTCAACTACTGGATTGCGAGCTCCCGGAGCGTGGACCTCCAGACACGGAT
CATGGAGCTGGAAGGCAGGGTCCGCAGGGCGGCTGCAGAGAGAGGCGCCGTGGAGCTGAAGA
AGAACGAGTTCCAGGGAGAGCTGGAGAAGCAGCGGGAGCAGCTTGACAAAATCCAGTCCAGC
CACAACTTCCAGCTGGAGAGCGTCAACAAGCTGTACCAGGACGAAAAGGCGGTTTTGGTGAA
TAACATCACCACAGGTGAGAGGCTCATCCGAGTGCTGCAAGACCAGTTAAAGACCCTGCAGA
GGAATTACGGCAGGCTGCAGCAGGATGTCCTCCAGTTTCAGAAGAACCAGACCAACCTGGAG
AGGAAGTTCTCCTACGACCTGAGCCAGTGCATCAATCAGATGAAGGAGGTGAAGGAACAGTG
TGAGGAGCGAATAGAAGAGGTCACCAAAAAGGGGAATGAAGCTGTAGCTTCCAGAGACCTGA
GTGAAAACAACGACCAGAGACAGCAGCTCCAAGCCCTCAGTGAGCCTCAGCCCAGGCTGCAG
GCAGCAGGCCTGCCACACACAGAGGTGCCACAAGGGAAGGGAAACGTGCTTGGTAACAGCAA
GTCCCAGACACCAGCCCCCAGTTCCGAAGTGGTTTTGGATTCAAAGAGACAAGTTGAGAAAG
AGGAAACCAATGAGATCCAGGTGGTGAATGAGGAGCCTCAGAGGGACAGGCTGCCGCAGGAG
CCAGGCCGGGAGCAGGTGGTGGAAGACAGACCTGTAGGTGGAAGAGGCTTCGGGGGAGCCGG
AGAACTGGGCCAGACCCCACAGGTGCAGGCTGCCCTGTCAGTGAGCCAGGAAAATCCAGAGA
TGGAGGGCCCTGAGCGAGACCAGCTTGTCATCCCCGACGGACAGGAGGAGGAGCAGGAAGCT
GCCGGGGAAGGGAGAAACCAGCAGAAACTGAGAGGAGAAGATGACTACAACATGGATGAAAA
TGAAGCAGAATCTGAGACAGACAAGCAAGCAGCCCTGGCAGGGAATGACAGAAACATAGATG
TTTTTAATGTTGAAGATCAGAAAAGAGACACCATAAATTTACTTGATCAGCGTGAAAAGCGG
AATCATACACTCTGAATTGAACTGGAATCACATATTTCACAACAGGGCCGAAGAGATGACTA
TAAAATGTTCATGAGGGACTGAATACTGAAAACTGTGAAATGTACTAAATAAAATGTACATCTGA
```

FIGURE 58

MMGLGNGRRSMKSPPLVLAALVACIIVLGFNYWIASSRSVDLQTRIMELEGRVRRAAAERGA

VELKKNEFQGELEKQREQLDKIQSSHNFQLESVNKLYQDEKAVLVNNITTGERLIRVLQDQL

KTLQRNYGRLQQDVLQFQKNQTNLERKFSYDLSQCINQMKEVKEQCEERIEEVTKKGNEAVA

SRDLSENNDQRQQLQALSEPQPRLQAAGLPHTEVPQGKGNVLGNSKSQTPAPSSEVVLDSKR

QVEKEETNEIQVVNEEPQRDRLPQEPGREQVVEDRPVGGRGFGGAGELGQTPQVQAALSVSQ

ENPEMEGPERDQLVIPDGQEEEQEAAGEGRNQQKLRGEDDYNMDENEAESETDKQAALAGND

RNIDVFNVEDQKRDTINLLDQREKRNHTL

Signal peptide:
amino acids 1-29

58/270

FIGURE 59

```
GGATGCAGAAAGCCTCAGTGTTGCTCTTCCTGGCCTGGGTCTGCTTCCTCTTCTACGCTGGCATTGCCCTCTTCA
CCAGTGGCTTCCTGCTCACCCGTTTGGAGCTCACCAACCATAGCAGCTGCCAAGAGCCCCCAGGCCCTGGGTCCC
TGCCATGGGGGAGCCAAGGGAAACCTGGGGCCTGCTGGATGGCTTCCCGATTTTCGCGGGTTGTGTTGGTGCTGA
TAGATGCTCTGCGATTTGACTTCGCCCAGCCCCAGCATTCACACGTGCCTAGAGAGCCTCCTGTCTCCCTACCCT
TCCTGGGCAAACTAAGCTCCTTGCAGAGGATCCTGGAGATTCAGCCCCACCATGCCCGGCTCTACCGATCTCAGG
TTGACCCTCCTACCACCACCATGCAGCGCCTCAAGGCCCTCACCACTGGCTCACTGCCTACCTTTATTGATGCTG
GTAGTAACTTCGCCAGCCACGCCATAGTGGAAGACAATCTCATTAAGCAGCTCACCAGTGCAGGAAGGCGTGTAG
TCTTCATGGGAGATGATACCTGGAAAGACCTTTTCCCTGGTGCTTTCTCCAAAGCTTTCTTCTTCCCATCCTTCA
ATGTCAGAGACCTAGACACAGTGGACAATGGCATCCTGGAACACCTCTACCCCACCATGGACAGTGGTGAATGGG
ACGTGCTGATTGCTCACTTCCTGGGTGTGGACCACTGTGGCCACAAGCATGGCCCTCACCACCCTGAAATGGCCA
AGAAACTTAGCCAGATGGACCAGGTGATCCAGGGACTTGTGGAGCGTCTGGAGAATGACACACTGCTGGTAGTGG
CTGGGGACCATGGGATGACCACAAATGGAGACCATGGAGGGGACAGTGAGCTGGAGGTCTCAGCTGCTCTCTTTC
TGTATAGCCCCACAGCAGTCTTCCCCAGCACCCCACCAGAGGAGCCAGAGGTGATTCCTCAAGTTAGCCTTGTGC
CCACGCTGGCCCTGCTGCTGGGCCTGCCCATCCCATTTGGGAATATCGGGGAAGTGATGGCTGAGCTATTCTCAG
GGGGTGAGGACTCCCAGCCCCACTCCTCTGCTTTAGCCCAAGCCTCAGCTCTCCATCTCAATGCTCAGCAGGTGT
CCCGATTTCTTCATACCTACTCAGCTGCTACTCAGGACCTTCAAGCTAAGGAGCTTCATCAGCTGCCAGAACCTCT
TCTCCAAGGCCTCTGCTGACTACCAGTGGCTTCTCCAGAGCCCCAAGGGGGCTGAGGCGACACTGCCGACTGTGA
TTGCTGAGCTGCAGCAGTTCCTGCGGGGAGCTCGGGCCATGTGCATCGAGTCTTGGGCTCGTTTCTCTCTGGTCC
GCATGGCGGGGGGTACTGCTCTCTTGGCTGCTTCCTGCTTTATCTGCCTGCTGGCATCTCAGTGGGCAATATCCC
CAGGCTTTCCATTCTGCCCTCTACTCCTGACACCTGTGGCCTGGGGCCTGGTTGGGGCCATAGCGTATGCTGGAC
TCCTGGGAACTATTGAGCTGAAGCTAGATCTAGTGCTTCTAGGGGCTGTGGCTGCAGTGAGCTCATTCCTCCCTT
TTCTGTGGAAAGCCTGGGCTGGCTGGGGGTCCAAGAGGCCCCTGGCAACCCTGTTTCCCATCCCTGGGCCCGTCC
TGTTACTCCTGCTGTTTCGCTTGGCTGTGTTCTTCTCTGATAGTTTTGTTGTAGCTGAGGCCAGGGCCACCCCCT
TCCTTTTGGGCTCATTCATCCTGCTCCTGGTTGTCCAGCTTCACTGGGAGGGCCAGCTGCTTCCACCTAAGCTAC
TCACAATGCCCCGCCTTGGCACTTCAGCCACAACAAACCCCCCACGGCACAATGGTGCATATGCCCTGAGGCTTG
GAATTGGGTTGCTTTTATGTACAAGGCTAGCTGGGCTTTTTCATCGTTGCCCTGAAGAGACACCTGTTTGCCACT
CCTCTCCCTGGCTGAGTCCTCTGGCATCCATGGTGGGTGGTCGAGCCAAGAATTTATGGTATGGAGCTTGTGTGG
CGGCGCTGGTGGCCCTGTTAGCTGCCGTGCGCTTGTGGCTTCGCCGCTATGGTAATCTCAAGAGCCCCGAGCCAC
CCATGCTCTTTGTGCGCTGGGGACTGCCCCTAATGGCATTGGGTACTGCTGCCTACTGGGCATTGGCGTCGGGGG
CAGATGAGGCTCCCCCCCCGTCTCCGGGTCCTGGTCTCTGGGGCATCCATGGTGCTGCCTCGGGCTGTAGCAGGGC
TGGCTGCTTCAGGGCTCGCGCTGCTGCTCTGGAAGCCTGTGACAGTGCTGGTGAAGGCTGGGGCAGGCGCTCCAA
GGACCAGGACTGTCCTCACTCCCTTCTCAGGCCCCCCACTTCTCAAGCTGACTTGGATTATGTGGTCCCTCAAA
TCTACCGACACATGCAGGAGGAGTTCCGGGGCCGGTTAGAGAGGACCAAATCTCAGGGTCCCCTGACTGTGGCTG
CTTATCAGTTGGGGAGTGTCTACTCAGCTGCTATGGTCACAGCCCTCACCCTGTTGGCCTTCCCACTTCTGCTGT
TGCATGCGGAGCGCATCAGCCTTGTGTTCCTGCTTCTGTTTCTGCAGAGCTTCCTTCTCCTACATCTGCTTGCTG
CTGGGATACCCGTCACCACCCCTGGTCCTTTTACTGTGCCATGGCAGGCAGTCTCGGCTTGGGCCCTCATGGCCA
CACAGACCTTCTACTCCACAGGCCACCAGCCTGTCTTTCCAGCCATCCATTGGCATGCAGCCTTCGTGGGATTCC
CAGAGGGTCATGGCTCCTGTACTTGGCTGCCTGCTTTGCTAGTGGGGAGCCAACACCTTTGCCTCCCACCTCCTCT
TTGCAGTAGGTTGCCCACTGCTCCTGCTCTGGCCTTTCCTGTGTGAGAGTCAAGGGCTGCGGAAGAGACAGCAGC
CCCCAGGGAATGAAGCTGATGCCAGAGTCAGACCCGAGGAGGAAGAGGAGCCACTGATGGAGATGCGGCTCCGGG
ATGCGCCTCAGCACTTCTATGCAGCACTGCTGCAGCTGGGCCTCAAGTACCTCTTTATCCTTGGTATTCAGATTC
TGGCCTGTGCCTTGGCAGCCTCCATCCTTCGCAGGCATCTCATGGTCTGGAAAGTGTTTGCCCCTAAGTTCATAT
TTGAGGCTGTGGGCTTCATTGTGAGCAGCGTGGGACTTCTCCTGGGCATAGCTTTGGTGATGAGAGTGGATGGTG
CTGTGAGCTCCTGGTTCAGGCAGCTATTTCTGGCCCAGCAGAGGTAGCCTAGTCTGTGATTACTGGCACTTGGCT
ACAGAGAGTGCTGGAGAACAGTGTAGCCTGGCCTGTACAGGTACTGGATGATCTGCAAGACAGGCTCAGCCATAC
TCTTACTATCATGCAGCCAGGGGCCGCTGACATCTAGGACTTCATTATTCTATAATTCAGGACCACAGTGGAGTA
TGATCCCTAACTCCTGATTTGGATGCATCTGAGGGACAAGGGGGCGGTCTCCGAAGTGGAATAAAATAGGCCGG
GCGTGGTGACTTGCACCTATAATCCCAGCACTTTGGGAGGCAGAGGTGGGAGGATTGCTTGGTCCCAGGAGTTCA
AGACCAGCCTGTGGAACATAACAAGACCCCGTCTCTACTATTTAAAAAAAAAGTGTAATAAAATGATAATAT
```

FIGURE 60

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62809

<subunit 1 of 1, 1089 aa, 1 stop

<MW: 118699, pI: 8.49, NX(S/T): 2

MQKASVLLFLAWVCFLFYAGIALFTSGFLLTRLELTNHSSCQEPPGPGSLPWGSQGKPGACW
MASRFSRVVLVLIDALRFDFAQPQHSHVPREPPVSLPFLGKLSSLQRILEIQPHHARLYRSQ
VDPPTTTMQRLKALTTGSLPTFIDAGSNFASHAIVEDNLIKQLTSAGRRVVFMGDDTWKDLF
PGAFSKAFFFPSFNVRDLDTVDNGILEHLYPTMDSGEWDVLIAHFLGVDHCGHKHGPHHPEM
AKKLSQMDQVIQGLVERLENDTLLVVAGDHGMTTNGDHGGDSELEVSAALFLYSPTAVFPST
PPEEPEVIPQVSLVPTLALLLGLPIPFGNIGEVMAELFSGGEDSQPHSSALAQASALHLNAQ
QVSRFLHTYSAATQDLQAKELHQLQNLFSKASADYQWLLQSPKGAEATLPTVIAELQQFLRG
ARAMCIESWARFSLVRMAGGTALLAASCFICLLASQWAISPGFPFCPLLLTPVAWGLVGAIA
YAGLLGTIELKLDLVLLGAVAAVSSFLPFLWKAWAGWGSKRPLATLFPIPGPVLLLLLFRLA
VFFSDSFVVAEARATPFLLGSFILLLVVQLHWEGQLLPPKLLTMPRLGTSATTNPPRHNGAY
ALRLGIGLLLCTRLAGLFHRCPEETPVCHSSPWLSPLASMVGGRAKNLWYGACVAALVALLA
AVRLWLRRYGNLKSPEPPMLFVRWGLPLMALGTAAYWALASGADEAPPRLRVLVSGASMVLP
RAVAGLAASGLALLLWKPVTVLVKAGAGAPRTRTVLTPFSGPPTSQADLDYVVPQIYRHMQE
EFRGRLERTKSQGPLTVAAYQLGSVYSAAMVTALTLLAFPLLLLHAERISLVFLLLFLQSFL
LLHLLAAGIPVTTPGPFTVPWQAVSAWALMATQTFYSTGHQPVFPAIHWHAAFVGFPEGHGS
CTWLPALLVGANTFASHLLFAVGCPLLLLWPFLCESQGLRKRQQPPGNEADARVRPEEEEEP
LMEMRLRDAPQHFYAALLQLGLKYLFILGIQILACALAASILRRHLMVWKVFAPKFIFEAVG
FIVSSVGLLLGIALVMRVDGAVSSWFRQLFLAQQR

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domains:

amino acids 317-341, 451-470, 481-500, 510-527, 538-555, 831-850,
1016-1034, 1052-1070

Leucine zipper pattern.

amino acids 843-864

N-glycosylation sites.

amino acids 37-40, 268-271

60/270

FIGURE 61

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGT

GTCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA<u>ATGT</u>

CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATT

TCATCACAAATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGT

CCTGACAGCTCCAGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAA

TATACTCCAATCTGAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCC

CAGTGTGTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGT

ACACGTGGAGTCCTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTG

CCAGGACTTTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTG

CCCATATCTATTACCGTGTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCA

CGTTGGCAAAGAGAAACACCCAGCAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAA

GATTCTTTGTGCCTGCTGAAAAAATCGTGATTAACTTTATCACCCTCAATATCTCGGATGAT

TCTAAAATTTCTCATCAGGATATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAA

TGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAG

GGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCT

CTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAAACAGTCATTGAATATGA

ATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGG

AGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCG

CAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCCTGGCGCAGGAGCA

CACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGATCCCC

AAACTGGCAGGCTGTGTATTCCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAG

CCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGC

TCCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGT

TATATGTGCAGATGGAAAAC<u>TGA</u>TGCCAACACTTCCTTTTGCCTTTTGTTTCCTGTGCAAAC

AAGTGAGTCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGT

TTGTCAGTGTCTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTG

GTTCATGCATGTAGGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGT

TGTTCTATGCAGAGAAAGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGG

TGGGTGT

FIGURE 62

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62815

<subunit 1 of 1, 442 aa, 1 stop

<MW: 49932, pI: 4.55, NX(S/T): 5

MSYNGLHQRVFKELKLLTLCSISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQ

QIYSNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQ

CARTLKDQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIYGNEFD

KRFFVPAEKIVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKH

LGYASHLMEIFCDSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSL

QEEVSTQGTLLESQAALAVLGPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWD

PQTGRLCIPSLSSFDQDSEGCEPSEGDGLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEW

GLYVQMEN

Important features:
Signal peptide:
amino acids 1-28

Transmembrane domain:
amino acids 140-163

N-glycosylation sites.
amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 63

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTCTCTGCGGGGAGACGCCAGCCTGCG

TCTGCC**ATG**GGGCTCGGGTTGAGGGGCTGGGGACGTCCTCTGCTGACTGTGGCCACCGCCCT

GATGCTGCCCGTGAAGCCCCCCGCAGGCTCCTGGGGGGCCCAGATCATCGGGGGCCACGAGG

TGACCCCCCACTCCAGGCCCTACATGGCATCCGTGCGCTTCGGGGGCCAACATCACTGCGGA

GGCTTCCTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCCACAGAGACCT

CCGCACTGGCCTGGTGGTGCTGGGCGCCCACGTCCTGAGTACTGCGGAGCCCACCCAGCAGG

TGTTTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCCAACGAC

ATCTGCCTGCTGCGGCTGAACGGCTCTGCTGTCCTGGGCCCTGCAGTGGGGCTGCTGAGGCT

GCCAGGGAGAAGGGCCAGGCCCCCCACAGCGGGGACACGGTGCCGGGTGGCTGGCTGGGGCT

TCGTGTCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGGCCAAGGTCCGAGTGCTGGAC

CCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATGCTCTGCACCCGCAG

TGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGGCCCCTGGTGTGCAGGA

ACCGGGCTCACGGCCTCGTTTCCTTCTCGGGCCTCTGGTGCGGCGACCCCAAGACCCCCGAC

GTGTACACGCAGGTGTCCGCCTTTGTGGCCTGGATCTGGGACGTGGTTCGGCGGAGCAGTCC

CCAGCCCGGCCCCCTGCCTGGGACCACCAGGCCCCCAGGAGAAGCCGCC**TGA**GCCACAACCT

TGCGGCATGCAAATGAGATGGCCGCTCCAGGCCTGGAATGTTCCGTGGCTGGGCCCCACGGG

AAGCCTGATGTTCAGGGTTGGGGTGGGACGGGCAGCGGTGGGGCACACCCATTCCACATGCA

AAGGGCAGAAGCAAACCCAGTAAAATGTTAACTGACAAAAAAAAAAAAAAAAAAAAGAAA

FIGURE 64

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62845

><subunit 1 of 1, 283 aa, 1 stop

><MW: 30350, pI: 9.66, NX(S/T): 2

MGLGLRGWGRPLLTVATALMLPVKPPAGSWGAQIIGGHEVTPHSRPYMASVRFGGQHHCGGF

LLRARWVVSAAHCFSHRDLRTGLVVLGAEVLSTAEPTQQVFGIDALTTHPDYHPMTHANDIC

LLRLNGSAVLGPAVGLLRLPGRRARPPTAGTRCRVAGWGFVSDFEELPPGLMEAKVRVLDPD

VCNSSWKGHLTLTMLCTRSGDSHRRGFCSADSGGPLVCRNRAHGLVSFSGLWCGDPKTPDVY

TQVSAFVAWIWDVVRRSSPQPGPLPGTTRPPGEAA

 Signal peptide:

amino acids 1-30

FIGURE 65

GAGCTACCCAGGCGGCTGGTGTGCAGCAAGCTCCGCGCCGACTCCGGACGCCTGACGCCTGA
CGCCTGTCCCCGGCCCGGC**ATG**AGCCGCTACCTGCTGCCGCTGTCGGCGCTGGGCACGGTAG
CAGGCGCCGCCGTGCTGCTCAAGGACTATGTCACCGGTGGGGCTTGCCCCAGCAAGGCCACC
ATCCCTGGGAAGACGGTCATCGTGACGGGCGCCAACACAGGCATCGGGAAGCAGACCGCCTT
GGAACTGGCCAGGAGAGGAGGCAACATCATCCTGGCCTGCCGAGACATGGAGAAGTGTGAGG
CGGCAGCAAAGGACATCCGCGGGGAGACCCTCAATCACCATGTCAACGCCCGGCACCTGGAC
TTGGCTTCCCTCAAGTCTATCCGAGAGTTTGCAGCAAAGATCATTGAAGAGGAGGAGCGAGT
GGACATTCTAATCAACAACGCGGGTGTGATGCGGTGCCCCCACTGGACCACCGAGGACGGCT
TCGAGATGCAGTTTGGCGTTAACCACCTGGGTCACTTTCTCTTGACAAACTTGCTGCTGGAC
AAGCTGAAAGCCTCAGCCCCTTCGCGGATCATCAACCTCTCGTCCCTGGCCCATGTTGCTGG
GCACATAGACTTTGACGACTTGAACTGGCAGACGAGGAAGTATAACACCAAAGCCGCCTACT
GCCAGAGCAAGCTCGCCATCGTCCTCTTCACCAAGGAGCTGAGCCGGCGGCTGCAAGGCTCT
GGTGTGACTGTCAACGCCCTGCACCCCGGCGTGGCCAGGACAGAGCTGGGCAGACACACGGG
CATCCATGGCTCCACCTTCTCCAGCACCACACTCGGGCCCATCTTCTGGCTGCTGGTCAAGA
GCCCCGAGCTGGCCGCCCAGCCCAGCACATACCTGGCCGTGGCGGAGGAACTGGCGGATGTT
TCCGGAAAGTACTTCGATGGACTCAAACAGAAGGCCCCGGCCCCCGAGGCTGAGGATGAGGA
GGTGGCCCGGAGGCTTTGGGCTGAAAGTGCCCGCCTGGTGGGCTTAGAGGCTCCCTCTGTGA
GGGAGCAGCCCCTCCCCAGA**TAA**CCTCTGGAGCAGATTTGAAAGCCAGGATGGCGCCTCCAG
ACCGAGGACAGCTGTCCGCCATGCCCGCAGCTTCCTGGCACTACCTGAGCCGGGAGACCCAG
GACTGGCGGCCGCCATGCCCGCAGTAGGTTCTAGGGGGCGGTGCTGGCCGCAGTGGACTGGC
CTGCAGGTGAGCACTGCCCCGGGCTCTGGCTGGTTCCGTCTGCTCTGCTGCCAGCAGGGGAG
AGGGGCCATCTGATGCTTCCCCTGGGAATCTAAACTGGGAATGGCCGAGGAGGAAGGGGCTC
TGTGCACTTGCAGGCCACGTCAGGAGAGCCAGCGGTGCCTGTCGGGGAGGGTTCCAAGGTGC
TCCGTGAAGAGCATGGGCAAGTTGTCTGACACTTGGTGGATTCTTGGGTCCCTGTGGGACCT
TGTGCATGCATGGTCCTCTCTGAGCCTTGGTTTCTTCAGCAGTGAGATGCTCAGAATAACTG
CTGTCTCCCATGATGGTGTGGTACAGCGAGCTGTTGTCTGGCTATGGCATGGCTGTGCCGGG
GGTGTTTGCTGAGGGCTTCCTGTGCCAGAGCCCAGCCAGAGAGCAGGTGCAGGTGTCATCCC
GAGTTCAGGCTCTGCACGGCATGGAGTGGGAACCCCACCAGCTGCTGCTACAGGACCTGGGA
TTGCCTGGGACTCCCACCTTCCTATCAATTCTCATGGTAGTCCAAACTGCAGACTCTCAAAC
TTGCTCATTT

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64842

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35932, pI: 8.45, NX(S/T): 1

MSRYLLPLSALGTVAGAAVLLKDYVTGGACPSKATIPGKTVIVTGANTGIGKQTALELARRG

GNIILACRDMEKCEAAAKDIRGETLNHHVNARHLDLASLKSIREFAAKIIEEEERVDILINN

AGVMRCPHWTTEDGFEMQFGVNHLGHFLLTNLLLDKLKASAPSRIINLSSLAHVAGHIDFDD

LNWQTRKYNTKAAYCQSKLAIVLFTKELSRRLQGSGVTVNALHPGVARTELGRHTGIHGSTF

SSTTLGPIFWLLVKSPELAAQPSTYLAVAEELADVSGKYFDGLKQKAPAPEAEDEEVARRLW

AESARLVGLEAPSVREQPLPR

Signal peptide:

amino acids 1-17

FIGURE 67

GAAGTTCGCGAGCGCTGGC**ATG**TGGTCCTGGGGCGCGGCTGGCGGCGCTGCTGGCGGTGCTG
GCGCTCGGGACAGGAGACCCAGAAAGGGCTGCGGCTCGGGGCGACACGTTCTCGGCGCTGAC
CAGCGTGGCGCGCGCCCTGGCGCCCGAGCGCCGGCTGCTGGGGCTGCTGAGGCGGTACCTGC
GCGGGGAGGAGGCGCGGCTGCGGGACCTGACTAGATTCTACGACAAGGTACTTTCTTTGCAT
GAGGATTCAACAACCCCTGTGGCTAACCCTCTGCTTGCATTTACTCTCATCAAACGCCTGCA
GTCTGACTGGAGGAATGTGGTACATAGTCTGGAGGCCAGTGAGAACATCCGAGCTCTGAAGG
ATGGCTATGAGAAGGTGGAGCAAGACCTTCCAGCCTTTGAGGACCTTGAGGGAGCAGCAAGG
GCCCTGATGCGGCTGCAGGACGTGTACATGCTCAATGTGAAAGGCCTGGCCCGAGGTGTCTT
TCAGAGAGTCACTGGCTCTGCCATCACTGACCTGTACAGCCCCAAACGGCTCTTTTCTCTCA
CAGGGGATGACTGCTTCCAAGTTGGCAAGGTGGCCTATGACATGGGGGATTATTACCATGCC
ATTCCATGGCTGGAGGAGGCTGTCAGTCTCTTCCGAGGATCTTACGGAGAGTGGAAGACAGA
GGATGAGGCAAGTCTAGAAGATGCCTTGGATCACTTGGCCTTTGCTTATTTCCGGGCAGGAA
ATGTTTCGTGTGCCCTCAGCCTCTCTCGGGAGTTTCTTCTCTACAGCCCAGATAATAAGAGG
ATGGCCAGGAATGTCTTGAAATATGAAAGGCTCTTGGCAGAGAGCCCCAACCACGTGGTAGC
TGAGGCTGTCATCCAGAGGCCCAATATACCCCACCTGCAGACCAGAGACACCTACGAGGGGC
TATGTCAGACCCTGGGTTCCCAGCCCACTCTCTACCAGATCCCTAGCCTCTACTGTTCCTAT
GAGACCAATTCCAACGCCTACCTGCTGCTCCAGCCCATCCGGAAGGAGGTCATCCACCTGGA
GCCCTACATTGCTCTCTACCATGACTTCGTCAGTGACTCAGAGGCTCAGAAAATTAGAGAAC
TTGCAGAACCATGGCTACAGAGGTCAGTGGTGGCATCAGGGGAGAAGCAGTTACAAGTGGAG
TACCGCATCAGCAAAAGTGCCTGGCTGAAGGACACTGTTGACCCAAAACTGGTGACCCTCAA
CCACCGCATTGCTGCCCTCACAGGCCTTGATGTCCGGCCTCCCTATGCAGAGTATCTGCAGG
TGGTGAACTATGGCATCGGAGGACACTATGAGCCTCACTTTGACCATGCTACGTCACCAAGC
AGCCCCCTCTACAGAATGAAGTCAGGAAACCGAGTTGCAACATTTATGATCTATCTGAGCTC
GGTGGAAGCTGGAGGAGCCACAGCCTTCATCTATGCCAACCTCAGCGTGCCTGTGGTTAGGA
ATGCAGCACTGTTTTGGTGGAACCTGCACAGGAGTGGTGAAGGGGACAGTGACACACTTCAT
GCTGGCTGTCCTGTCCTGGTGGGAGATAAGTGGGTGGCCAACAAGTGGATACATGAGTATGG
ACAGGAATTCCGCAGACCCTGCAGCTCCAGCCTGAAGAC**TGA**ACTGTTGGCAGAGAGAAGC
TGGTGGAGTCCTGTGGCTTTCCAGAGAAGCCAGGAGCCAAAAGCTGGGGTAGGAGAGGAGAA
AGCAGAGCAGCCTCCTGGAAGAAGGCCTTGTCAGCTTTGTCTGTGCCTCGCAAATCAGAGGC
AAGGGAGAGGTTGTTACCAGGGGACACTGAGAATGTACATTTGATCTGCCCCAGCCACGGAA
GTCAGAGTAGGATGCACAGTACAAAGGAGGGGGGAGTGGAGGCCTGAGAGGGAAGTTTCTGG
AGTTCAGATACTCTCTGTTGGGAACAGGACATCTCAACAGTCTCAGGTTCGATCAGTGGGTC
TTTTGGCACTTTGAACCTTGACCACAGGGACCAAGAAGTGGCAATGAGGACACCTGCAGGAG
GGGCTAGCCTGACTCCCAGAACTTTAAGACTTTCTCCCCACTGCCTTCTGCTGCAGCCCAAG
CAGGGAGTGTCCCCCTCCCAGAAGCATATCCCAGATGAGTGGTACATTATATAAGGATTTTT
TTTAAGTTGAAAACAACTTTCTTTTCTTTTTGTATGATGGTTTTTTAACACAGTCATTAAAA
ATGTTTATAAATCAAAA

FIGURE 68

MGPGARLAALLAVLALGTGDPERAAARGDTFSALTSVARALAPERRLLGLLRRYLRGEEARL
RDLTRFYDKVLSLHEDSTTPVANPLLAFTLIKRLQSDWRNVVHSLEASENIRALKDGYEKVE
QDLPAFEDLEGAARALMRLQDVYMLNVKGLARGVFQRVTGSAITDLYSPKRLFSLTGDDCFQ
VGKVAYDMGDYYHAIPWLEEAVSLFRGSYGEWKTEDEASLEDALDHLAFAYFRAGNVSCALS
LSREFLLYSPDNKRMARNVLKYERLLAESPNHVVAEAVIQRPNIPHLQTRDTYEGLCQTLGS
QPTLYQIPSLYCSYETNSNAYLLLQPIRKEVIHLEPYIALYHDFVSDSEAQKIRELAEPWLQ
RSVVASGEKQLQVEYRISKSAWLKDTVDPKLVTLNHRIAALTGLDVRPPYAEYLQVVNYGIG
GHYEPHFDHATSPSSPLYRMKSGNRVATFMIYLSSVEAGGATAFIYANLSVPVVRNAALFWW
NLHRSGEGDSDTLHAGCPVLVGDKWVANKWIHEYGQEFRRPCSSSPED

Signal peptide:
amino acids 1-19

FIGURE 69

GAGATAGGGAGTCTGGGTTTAAGTTCCTGCTCCATCTCAGGAGCCCCTGCTCCCACCCCTAG

GAAGCCACCAGACTCCACGGTGTGGGGCCAATCAGGTGGAATCGGCCCTGGCAGGTGGGGCC

ACGAGCGCTGGCTGAGGGACCGAGCCGGAGAGCCCCGGAGCCCCCGTAACCCGCGCGGGGAG

CGCCCAGG<u>ATG</u>CCGCGCGGGGACTCGGAGCAGGTGCGCTACTGCGCGCGCTTCTCCTACCTC

TGGCTCAAGTTTTCACTTATCATCTATTCCACCGTGTTC<u>T</u>GGCTGATTGGGGCCCTGGTCCT

GTCTGTGGGCATCTATGCAGAGGTTGAGCGGCAGAAATATAAAACCCTTGAAAGTGCCTTCC

TGGCTCCAGCCATCATCCTCATCCTCCTGGGCGTCGTCATGTTCATGGTCTCCTTCATTGGT

GTGCTGGCGTCCCTCCGTGACAACCTGTACCTTCTCCAAGCATTCATGTACATCCTTGGGAT

CTGCCTCATCATGGAGCTCATTGGTGGCGTGGTGGCCTTGACCTTCCGGAACCAGACCATTG

ACTTCCTGAACGACAACATTCGAAGAGGAATTGAGAACTACTATGATGATCTGGACTTCAAA

AACATCATGGACTTTGTTCAGAAAAAGTTCAAGTGCTGTGGCGGGGAGGACTACCGAGATTG

GAGCAAGAATCAGTACCACGACTGCAGTGCCCCTGGACCCCTGGCCTGTGGGGTGCCCTACA

CCTGCTGCATCAGGAACACGACAGAAGTTGTCAACACCATGTGTGGCTACAAAACTATCGAC

AAGGAGCGTTTCAGTGTGCAGGATGTCATCTACGTGCGGGGCTGCACCAACGCCGTGATCAT

CTGGTTCATGGACAACTACACCATCATGGCGTGCATCCTCCTGGGCATCCTGCTTCCCCAGT

TCCTGGGGGTGCTGCTGACGCTGCTGTACATCACCCGGGTGGAGGACATCATCATGGAGCAC

TCTGTCACTGATGGGCTCCTGGGGCCCGGTGCCAAGCCCAGCGTGGAGGCGGCAGGCACGGG

ATGCTGCTTGTGCTACCCCAAT<u>TAG</u>GGCCCAGCCTGCCATGGCAGCTCCAACAAGGACCGTC

TGGGATAGCACCTCTCAGTCAACATCGTGGGGCTGGACAGGGCTGCGGCCCCTCTGCCCACA

CTCAGTACTGACCAAAGCCAGGGCTGTGTGTGCCTGTGTGTAGGTCCCACGGCCTCTGCCTC

CCCAGGGAGCAGAGCCTGGGCCTCCCCTAAGAGGCTTTCCCCGAGGCAGCTCTGGAATCTGT

GCCCACCTGGGGCCTGGGGAACAAGGCCCTCCTTTCTCCAGGCCTGGGCTACAGGGGAGGGA

GAGCCTGAGGCTCTGCTCAGGGCCCATTTCATCTCTGGCAGTGCCTTGGCGGTGGTATTCAA

GGCAGTTTTGTAGCACCTGTAATTGGGGAGAGGGAGTGTGCCCCTCGGGGCAGGAGGGAAGG

GCATCTGGGGAAGGGCAGGAGGGAAGAGCTGTCCATGCAGCCACGCCCATGGCCAGGTTGGC

CTCTTCTCAGCCTCCCAGGTGCCTTGAGCCCTCTTGCAAGGGCGGCTGCTTCCTTGAGCCTA

GTTTTTTTTTACGTGATTTTTGTAACATTCATTTTTTTGTACAGATAACAGGAGTTTCTGAC

TAATCAAAGCTGGTATTTCCCCGCATGTCTTATTCTTGCCCTTCCCCCAACCAGTTTGTTAA

TCAAACAATAAAAACATGTTTTGTTTTGTTTTTAAAAAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64863

><subunit 1 of 1, 294 aa, 1 stop

><MW: 33211, pI: 5.35, NX(S/T): 3

MPRGDSEQVRYCARFSYLWLKFSLIIYSTVFWLIGALVLSVGIYAEVERQKYKTLESAFLAP

AIILILLGVVMFMVSFIGVLASLRDNLYLLQAFMYILGICLIMELIGGVVALTFRNQTIDFL

NDNIRRGIENYYDDLDFKNIMDFVQKKFKCCGGEDYRDWSKNQYHDCSAPGPLACGVPYTCC

IRNTTEVVNTMCGYKTIDKERFSVQDVIYVRGCTNAVIIWFMDNYTIMACILLGILLPQFLG

VLLTLLYITRVEDIIMEHSVTDGLLGPGAKPSVEAAGTGCCLCYPN

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 22-42, 57-85, 93-116, 230-257

70/270

FIGURE 71

GAGGAGCGGGCCGAGGACTCCAGCGTGCCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGA

CACCTGGGAAGATGGCCGGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACC

TTGATCCAAGCCACCCTCAGTCCCACTGCAGTTCTCATCCTCGGCCCAAAAGTCATCAAAGA

AAAGCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGC

TCAGTGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACC

GTCCTGAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAA

GCCCTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCA

ACACGCCCCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATC

CGCATGGACACCAGTGCAAGTGGCCCCACCCGCCTGGTCCTCAGTGACTGTGCCACCAGCCA

TGGGAGCCTGCGCATCCAACTGCTGTATAAGCTCTCCTTCCTGGTGAACGCCTTAGCTAAGC

AGGTCATGAACCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTGTCCCGTG

ATCGAGGCTTCCTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCCATTTC

CCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATTC

AGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCAATAAC

TCTGCAGCTTCCCTGACAATGCCCACCCTGGACAACATCCCGTTCAGCCTCATCGTGAGTCA

GGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGAATTCATGGTCCTGTTGG

ACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAG

GCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTT

TTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTTTCCCTCCA

GTGAAGCCCTCCGCCCTTTGTTCACCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTAC

ACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGAT

GAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCC

ACTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTG

AAGGCCTTGGGATTCGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCC

AGCCTCCTTGTGGAAACCCAGCTCTCCTGTCTCCCAGTGAAGACTTGGATGGCAGCCATCAG

GGAAGGCTGGGTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAAT

CAATAAACACTTGCCTGTGAAAAA

FIGURE 72

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64881
><subunit 1 of 1, 484 aa, 1 stop
><MW: 52468, pI: 7.14, NX(S/T): 3

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLLSAM
REKPAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPL
VKTIVEFHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMN
LLVPSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYL
GAKLLDSQGKVTKWFNNSAASLTMPTLDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVL
PESAHRLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEAL
RPLFTLGIEASSEAQFYTKGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSIL
LPNQNGKLRSGVPVSLVKALGFEAAESSLTKDALVLTPASLWKPSSPVSQ

Important features of the protein:
Signal peptide:
amino acids 1-21

N-glycosylation sites.
amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.
amino acids 412-415

LBP / BPI / CETP family proteins.
amino acids 407-457

FIGURE 73

GAGCGAAC<u>ATGG</u>CAGCGCGTTGGCGGTTTTGGTGTGTCTCTGTGACCATGGTGGTGGCGCTG
CTCATCGTTTGCGACGTTCCCTCAGCCTCTGCCCAAAGAAAGAAGGAGATGGTGTTATCTGA
AAAGGTTAGTCAGCTGATGGAATGGACTAACAAAAGACCTGTAATAAGAATGAATGGAGACA
AGTTCCGTCGCCTTGTGAAAGCCCCACCGAGAAATTACTCCGTTATCGTCATGTTCACTGCT
CTCCAACTGCATAGACAGTGTGTCGTTTGCAAGCAAGCTGATGAAGAATTCCAGATCCTGGC
AAACTCCTGGCGATACTCCAGTGCATTCACCAACAGGATATTTTTTGCCATGGTGGATTTTG
ATGAAGGCTCTGATGTATTTCAGATGCTAAACATGAATTCAGCTCCAACTTTCATCAACTTT
CCTGCAAAAGGGAAACCCAAACGGGGTGATACATATGAGTTACAGGTGCGGGGTTTTTCAGC
TGAGCAGATTGCCCGGTGGATCGCCGACAGAACTGATGTCAATATTAGAGTGATTAGACCCC
CAAATTATGCTGGTCCCCTTATGTTGGGATTGCTTTTGGCTGTTATTGGTGGACTTGTGTAT
CTTCGAAGAAGTAATATGGAATTTCTCTTTAATAAAACTGGATGGGCTTTTGCAGCTTTGTG
TTTTGTGCTTGCTATGACATCTGGTCAAATGTGGAACCATATAAGAGGACCACCATATGCCC
ATAAGAATCCCCACACGGGACATGTGAATTATATCCATGGAAGCAGTCAAGCCCAGTTTGTA
GCTGAAACACACATTGTTCTTCTGTTTAATGGTGGAGTTACCTTAGGAATGGTGCTTTTATG
TGAAGCTGCTACCTCTGACATGGATATTGGAAAGCGAAAGATAATGTGTGTGGCTGGTATTG
GACTTGTTGTATTATTCTTCAGTTGGATGCTCTCTATTTTTAGATCTAAATATCATGGCTAC
CCATACAGCTTTCTGATGAGT<u>TAA</u>AAAGGTCCCAGAGATATATAGACACTGGAGTACTGGAA
ATTGAAAAACGAAAATCGTGTGTGTTTGAAAAGAAGAATGCAACTTGTATATTTTGTATTAC
CTCTTTTTTTCAAGTGATTTAAATAGTTAATCATTTAACCAAAGAAGATGTGTAGTGCCTTA
ACAAGCAATCCTCTGTCAAAATCTGAGGTATTTGAAAATAATTATCCTCTTAACCTTCTCTT
CCCAGTGAACTTTATGGAACATTTAATTTAGTACAATTAAGTATATTATAAAAATTGTAAAA
CTACTACTTTGTTTTAGTTAGAACAAAGCTCAAAACTACTTTAGTTAACTTGGTCATCTGAT
TTTATATTGCCTTATCCAAAGATGGGGAAAGTAAGTCCTGACCAGGTGTTCCCACATATGCC
TGTTACAGATAACTACATTAGGAATTCATTCTTAGCTTCTTCATCTTTGTGTGGATGTGTAT
ACTTTACGCATCTTTCCTTTTGAGTAGAGAAATTATGTGTGTCATGTGGTCTTCTGAAAATG
GAACACCATTCTTCAGAGCACACGTCTAGCCCTCAGCAAGACAGTTGTTTCTCCTCCTCCTT
GCATATTTCCTACTGCGCTCCAGCCTGAGTGATAGAGTGAGACTCTGTCTCAAAAAAAAGTA
TCTCTAAATACAGGATTATAATTTCTGCTTGAGTATGGTGTTAACTACCTTGTATTTAGAAA
GATTTCAGATTCATTCCATCTCCTTAGTTTTCTTTTAAGGTGACCCATCTGTGATAAAAATA
TAGCTTAGTGCTAAAATCAGTGTAACTTATACATGGCCTAAAATGTTTCTACAAATTAGAGT
TTGTCACTTATTCCATTTGTACCTAAGAGAAAAATAGGCTCAGTTAGAAAAGGACTCCCTGG
CCAGGCGCAGTGACTTACGCCTGTAATCTCAGCACTTTGGGAGGCCAAGGCAGGCAGATCAC
GAGGTCAGGAGTTCGAGACCATCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATAT
AAAAATTAGCTGGGTGTGGTGGCAGGAGCCTGTAATCCCAGCTACACAGGAGGCTGAGGCAC
GAGAATCACTTGAACTCAGGAGATGGAGGTTTCAGTGAGCCGAGATCACGCCACTGCACTCC
AGCCTGGCAACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAA

FIGURE 74

MAARWRFWCVSVTMVVALLIVCDVPSASAQRKKEMVLSEKVSQLMEWTNKRPVIRMNGDKFR

RLVKAPPRNYSVIVMFTALQLHRQCVVCKQADEEFQILANSWRYSSAFTNRIFFAMVDFDEG

SDVFQMLNMNSAPTFINFPAKGKPKRGDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNY

AGPLMLGLLLAVIGGLVYLRRSNMEFLFNKTGWAFAALCFVLAMTSGQMWNHIRGPPYAHKN

PHTGHVNYIHGSSQAQFVAETHIVLLFNGGVTLGMVLLCEAATSDMDIGKRKIMCVAGIGLV

VLFFSWMLSIFRSKYHGYPYSFLMS

Signal peptide:

amino acids 1-29

Transmembrane domains:

amino acids 183-205, 217-237, 217-287, 301-321

FIGURE 75

```
AAGCAACCAAACTGCAAGCTTTGGGAGTTGTTCGCTGTCCCTGCCCTGCTCTGCTAGGGAGA
GAACGCCAGAGGGAGGCGGCTGGCCCGGCGGCAGGCTCTCAGAACCGCTACCGGCGATGCTA
CTGCTGTGGGTGTCGGTGGTCGCAGCCTTGGCGCTGGCGGTACTGGCCCCCGGAGCAGGGGA
GCAGAGGCGGAGAGCAGCCAAAGCGCCCAATGTGGTGCTGGTCGTGAGCGACTCCTTCGATG
GAAGGTTAACATTTCATCCAGGAAGTCAGGTAGTGAAACTTCCTTTTATCAACTTTATGAAG
ACACGTGGGACTTCCTTTCTGAATGCCTACACAAACTCTCCAATTTGTTGCCCATCACGCGC
AGCAATGTGGAGTGGCCTCTTCACTCACTTAACAGAATCTTGGAATAATTTTAAGGGTCTAG
ATCCAAATTATACAACATGGATGGATGTCATGGAGAGGCATGGCTACCGAACACAGAAATTT
GGGAAACTGGACTATACTTCAGGACATCACTCCATTAGTAATCGTGTGGAAGCGTGGACAAG
AGATGTTGCTTTCTTACTCAGACAAGAAGGCAGGCCCATGGTTAATCTTATCCGTAACAGGA
CTAAAGTCAGAGTGATGGAAAGGGATTGGCAGAATACAGACAAAGCAGTAAACTGGTTAAGA
AAGGAAGCAATTAATTACACTGAACCATTTGTTATTTACTTGGGATTAAATTTACCACACCC
TTACCCTTCACCATCTTCTGGAGAAAATTTTGGATCTTCAACATTTCACACATCTCTTTATT
GGCTTGAAAAAGTGTCTCATGATGCCATCAAAATCCCAAAGTGGTCACCTTTGTCAGAAATG
CACCCTGTAGATTATTACTCTTCTTATACAAAAAACTGCACTGGAAGATTTACAAAAAAAGA
AATTAAGAATATTAGAGCATTTTATTATGCTATGTGTGCTGAGACAGATGCCATGCTTGGTG
AAATTATTTTGGCCCTTCATCAATTAGATCTTCTTCAGAAAACTATTGTCATATACTCCTCA
GACCATGGAGAGCTGGCCATGGAACATCGACAGTTTTATAAAAATGAGCATGTACGAGGCTAG
TGCACATGTTCCGCTTTTGATGATGGGACCAGGAATTAAAGCCGGCCTACAAGTATCAAATG
TGGTTTCTCTTGTGGATATTTACCCTACCATGCTTGATATTGCTGGAATTCCTCTGCCTCAG
AACCTGAGTGGATACTCTTTGTTGCCGTTATCATCAGAAACATTTAAGAATGAACATAAAGT
CAAAAACCTGCATCCACCCTGGATTCTGAGTGAATTCCATGGATGTAATGTGAATGCCTCCA
CCTACATGCTTCGAACTAACCACTGGAAATATATAGCCTATTCGGATGGTGCATCAATATTG
CCTCAACTCTTTGATCTTTCCTCGGATCCAGATGAATTAACAAATGTTGCTGTAAAATTTCC
AGAAATTACTTATTCTTTGGATCAGAAGCTTCATTCCATTATAAACTACCCTAAAGTTTCTG
CTTCTGTCCACCAGTATAATAAAGAGCAGTTTATCAAGTGGAAACAAAGTATAGGACAGAAT
TATTCAAACGTTATAGCAAATCTTAGGTGGCACCAAGACTGGCAGAAGGAACCAAGGAAGTA
TGAAAATGCAATTGATCAGTGGCTTAAAACCCATATGAATCCAAGAGCAGTTTGAACAAAAA
GTTTAAAAATAGTGTTCTAGAGATACATATAAATATATTACAAGATCATAATTATGTATTTT
AAATGAAACAGTTTTAATAATTACCAAGTTTTGGCCGGGCACAGTGGCTCACACCTGTAATC
CCAGGACTTTGGGAGGCTGAGGAAAGCAGATCACAAGGTCAAGAGATTGAGACCATCCTGGC
CAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTAGCTGGGCGCGGTGGTGCACA
CCTATAGTCTCAGCTACTCAGAGGCTGAGGCAGGAGGATCGCTTGAACCCGGGAGGCAGCAG
TTGCAGTGAGCTGAGATTGCGCCACTGTACTCCAGCCTGGCAACAGAGTGAGACTGTGTCGC
AAAAAAATAAAAATAAAATAATAATAATTACCAATTTTTCATTATTTTGTAAGAATGTAGTG
TATTTTAAGATAAAATGCCAATGATTATAAAATCACATATTTTCAAAAATGGTTATTATTTA
GGCCTTTGTACAATTTCTAACAATTTAGTGGAAGTATCAAAAGGATTGAAGCAAATACTGTA
ACAGTTATGTTCCTTTAAATAATAGAGAATATAAAATATTGTAATAATATGTATCATAAAAT
AGTTGTATGTGAGCATTTGATGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

FIGURE 76

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64885

<subunit 1 of 1, 536 aa, 1 stop

<MW: 61450, pI: 9.17, NX(S/T): 7

MLLLWVSVVAALALAVLAPGAGEQRRRAAKAPNVVLVVSDSFDGRLTFHPGSQVVKLPFINF

MKTRGTSFLNAYTNSPICCPSRAAMWSGLFTHLTESWNNFKGLDPNYTTWMDVMERHGYRTQ

KFGKLDYTSGHHSISNRVEAWTRDVAFLLRQEGRPMVNLIRNRTKVRVMERDWQNTDKAVNW

LRKEAINYTEPFVIYLGLNLPHPYPSPSSGENFGSSTFHTSLYWLEKVSHDAIKIPKWSPLS

EMHPVDYYSSYTKNCTGRFTKKEIKNIRAFYYAMCAETDAMLGEIILALHQLDLLQKTIVIY

SSDHGELAMEHRQFYKMSMYEASAHVPLLMMGPGIKAGLQVSNVVSLVDIYPTMLDIAGIPL

PQNLSGYSLLPLSSETFKNEHKVKNLHPPWILSEFHGCNVNASTYMLRTNHWKYIAYSDGAS

ILPQLFDLSSDPDELTNVAVKFPEITYSLDQKLHSIINYPKVSASVHQYNKEQFIKWKQSIG

QNYSNVIANLRWHQDWQKEPRKYENAIDQWLKTHMNPRAV

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-111, 166-169, 193-196, 262-265, 375-378, 413-416,

498-501

Sulfatases proteins:

amino acids 286-315, 359-369, 78-97

FIGURE 77

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAG

AGCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCC**ATG**

GCCTCTCTTGGCCTCCAACTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACT

GGTTGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAG

CAGTTGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAG

TGTGACATCTATAGCACCCTTCTGGGCCTGCCCGCTGACATCCAGGCTGCCCAGGCCATGAT

GGTGACATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCA

CAGTCTTCTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTC

ATCCTTGGAGGCCTCCTGGGATTCATTCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGA

CTTCTACTCACCACTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGG

GCATTATTTCTTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCCTGCTCATCC

CAGAGAAATCGCTCCAACTACTACGATGCCTACCAAGCCCAACCTCTTGCCACAAGGAGCTC

TCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCTACAGCCTGACAGGGT

ATGTG**TGA**AGAACCAGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAG

CACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGTGTCAGAAGGTGCTGCTGAGG

ATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATG

CAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCCTCACCTTGCTGCTC

CCCTGCCCTAAGTCCCCAACCCTCAACTTGAAACCCATTCCCTTAAGCCAGGACTCAGAGG

ATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCACTG

ACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGG

GGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCC

TCCAAAGAAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCC

AGACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAG

GATGCAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAAATA

FIGURE 78

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64886

><subunit 1 of 1, 230 aa, 1 stop

><MW: 24549, pI: 8.56, NX(S/T): 1

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGIT

QCDIYSTLLGLPADIQAAQAMMVTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVF

FILGGLLGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCS

SQRNRSNYYDAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

78/270

FIGURE 79

GCACTGCTGCTGTCCCATCAGCTGCTCTGAAGCTCC<u>ATG</u>GTGCCCAGAATCTTCGCTCCTGC

TTATGTGTCAGTCTGTCTCCTCCTCTTGTGTCCAAGGGAAGTCATCGCTCCCGCTGGCTCAG

AACCATGGCTGTGCCAGCCGGCACCCAGGTGTGGAGACAAGATCTACAACCCCTTGGAGCAG

TGCTGTTACAATGACGCCATCGTGTCCCTGAGCGAGACCCGCCAATGTGGTCCCCCCTGCAC

CTTCTGGCCCTGCTTTGAGCTCTGCTGTCTTGATTCCTTTGGCCTCACAAACGATTTTGTTG

TGAAGCTGAAGGTTCAGGGTGTGAATTCCCAGTGCCACTCATCTCCCATCTCCAGTAAATGT

GAAAGCAGAAGACGTTTTCCC<u>TGA</u>GAAGACATAGAAAGAAAATCAACTTTCACTAAGGCATC

TCAGAAACATAGGCTAAGGTAATATGTGTACCAGTAGAGAAGCCTGAGGAATTTACAAAATG

ATGCAGCTCCAAGCCATTGTATGGCCCATGTGGGAGACTGATGGGACATGGAGAATGACAGT

AGATTATCAGGAAATAAATAAAGTGGTTTTTCCAATGTACACACCTGTAAAA

FIGURE 80

MVPRIFAPAYVSVCLLLLCPREVIAPAGSEPWLCQPAPRCGDKIYNPLEQCCYNDAIVSLSE
TRQCGPPCTFWPCFELCCLDSFGLTNDFVVKLKVQGVNSQCHSSPISSKCESRRRFP

Signal peptide:
amino acids 1-25

FIGURE 81

CTCCACTGCAACCACCCAGAGCC<u>ATG</u>GCTCCCCGAGGCTGCATCGTAGCTGTCTTTGCCATT

TTCTGCATCTCCAGGCTCCTCTGCTCACACGGAGCCCCAGTGGCCCCCATGACTCCTTACCT

GATGCTGTGCCAGCCACACAAGAGATGTGGGGACAAGTTCTACGACCCCCTGCAGCACTGTT

GCTATGATGATGCCGTCGTGCCCTTGGCCAGGACCCAGACGTGTGGAAACTGCACCTTCAGA

GTCTGCTTTGAGCAGTGCTGCCCCTGGACCTTCATGGTGAAGCTGATAAACCAGAACTGCGA

CTCAGCCCGGACCTCGGATGACAGGCTTTGTCGCAGTGTCAGC<u>TAA</u>TGGAACATCAGGGGAA

CGATGACTCCTGGATTCTCCTTCCTGGGTGGGCCTGGAGAAAGAGGCTGGTGTTACCTGAGA

TCTGGGATGCTGAGTGGCTGTTTGGGGGCCAGAGAAACACACACTCAACTGCCCACTTCATT

CTGTGACCTGTCTGAGGCCCACCCTGCAGCTGCCCTGAGGAGGCCCACAGGTCCCCTTCTAG

AATTCTGGACAGCATGAGATGCGTGTGCTGATGGGGGCCCAGGGACTCTGAACCCTCCTGAT

GACCCCTATGGCCAACATCAACCCGGCACCACCCCAAGGCTGGCTGGGGAACCCTTCACCCT

TCTGTGAGATTTTCCATCATCTCAAGTTCTCTTCTATCCAGGAGCAAAGCACAGGATCATAA

TAAATTTATGTACTTTATAAATGAAAA

81 /270

FIGURE 82

MAPRGCIVAVFAIFCISRLLCSHGAPVAPMTPYLMLCQPHKRCGDKFYDPLQHCCYDDAVVP
LARTQTCGNCTFRVCFEQCCPWTFMVKLINQNCDSARTSDDRLCRSVS

Signal peptide:
amino acids 1-24

FIGURE 83

```
GGGGGCGGGTGCCTGGAGCACGGCGCTGGGGCCGCCCGCAGCGCTCACTCGCTCGCACTCAG
TCGCGGGAGGCTTCCCCGCGCCGGCCGCGTCCCGCCCGCTCCCCGGCACCAGAAGTTCCTCT
GCGCGTCCGACGGCGACATGGGCGTCCCCACGGCCCTGGAGGCCGGCAGCTGGCGCTGGGGA
TCCCTGCTCTTCGCTCTCTTCCTGGCTGCGTCCCTAGGTCCGGTGGCAGCCTTCAAGGTCGC
CACGCCGTATTCCCTGTATGTCTGTCCCGAGGGGCAGAACGTCACCCTCACCTGCAGGCTCT
TGGGCCCTGTGGACAAAGGGCACGATGTGACCTTCTACAAGACGTGGTACCGCAGCTCGAGG
GGCGAGGTGCAGACCTGCTCAGAGCGCCGGCCCATCCGCAACCTCACGTTCCAGGACCTTCA
CCTGCACCATGGAGGCCACCAGGCTGCCAACACCAGCCACGACCTGGCTCAGCGCCACGGGC
TGGAGTCGGCCTCCGACCACCATGGCAACTTCTCCATCACCATGCGCAACCTGACCCTGCTG
GATAGCGGCCTCTACTGCTGCCTGGTGGTGGAGATCAGGCACCACCACTCGGAGCACAGGGT
CCATGGTGCCATGGAGCTGCAGGTGCAGACAGGCAAAGATGCACCATCCAACTGTGTGGTGT
ACCCATCCTCCTCCCAGGATAGTGAAAACATCACGGCTGCAGCCCTGGCTACGGGTGCCTGC
ATCGTAGGAATCCTCTGCCTCCCCCTCATCCTGCTCCTGGTCTACAAGCAAAGGCAGGCAGC
CTCCAACCGCCGTGCCCAGGAGCTGGTGCGGATGGACAGCAACATTCAAGGGATTGAAAACC
CCGGCTTTGAAGCCTCACCACCTGCCCAGGGGATACCCGAGGCCAAAGTCAGGCACCCCCTG
TCCTATGTGGCCCAGCGGCAGCCTTCTGAGTCTGGGCGGCATCTGCTTTCGGAGCCCAGCAC
CCCCCTGTCTCCTCCAGGCCCCGGAGACGTCTTCTTCCCATCCCTGGACCCTGTCCCTGACT
CTCCAAACTTTGAGGTCATCTAGCCCAGCTGGGGGACAGTGGGCTGTTGTGGCTGGGTCTGG
GGCAGGTGCATTTGAGCCAGGGCTGGCTCTGTGAGTGGCCTCCTTGGCCTCGGCCCTGGTTC
CCTCCCTCCTGCTCTGGGCTCAGATACTGTGACATCCCAGAAGCCCAGCCCCTCAACCCCTC
TGGATGCTACATGGGGATGCTGGACGGCTCAGCCCCTGTTCCAAGGATTTTGGGGTGCTGAG
ATTCTCCCCTAGAGACCTGAAATTCACCAGCTACAGATGCCAAATGACTTACATCTTAAGAA
GTCTCAGAACGTCCAGCCCTTCAGCAGCTCTCGTTCTGAGACATGAGCCTTGGGATGTGGCA
GCATCAGTGGGACAAGATGGACACTGGGCCACCCTCCCAGGCACCAGACACAGGGCACGGTG
GAGAGACTTCTCCCCCGTGGCCGCCTTGGCTCCCCGTTTTGCCCGAGGCTGCTCTTCTGTC
AGACTTCCTCTTTGTACCACAGTGGCTCTGGGGCCAGGCCTGCCTGCCCACTGGCCATCGCC
ACCTTCCCCAGCTGCCTCCTACCAGCAGTTTCTCTGAAGATCTGTCAACAGGTTAAGTCAAT
CTGGGGCTTCCACTGCCTGCATTCCAGTCCCCAGAGCTTGGTGGTCCCGAAACGGGAAGTAC
ATATTGGGGCATGGTGGCCTCCGTGAGCAAATGGTGTCTTGGGCAATCTGAGGCCAGGACAG
ATGTTGCCCCACCCACTGGAGATGGTGCTGAGGGAGGTGGGTGGGGCCTTCTGGGAAGGTGA
GTGGAGAGGGGCACCTGCCCCCCGCCCTCCCCATCCCCTACTCCCACTGCTCAGCGCGGGCC
ATTGCAAGGGTGCCACACAATGTCTTGTCCACCCTGGGACACTTCTGAGTATGAAGCGGGAT
GCTATTAAAAACTACATGGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGA
```

83/270

FIGURE 84

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64897
><subunit 1 of 1, 311 aa, 1 stop
><MW: 33908, pI: 6.87, NX(S/T): 6
MGVPTALEAGSWRWGSLLFALFLAASLGPVAAFKVATPYSLYVCPEGQNVTLTCRLLGPVDK
GHDVTFYKTWYRSSRGEVQTCSERRPIRNLTFQDLHLHHGGHQAANTSHDLAQRHGLESASD
HHGNFSITMRNLTLLDSGLYCCLVVEIRHHHSEHRVHGAMELQVQTGKDAPSNCVVYPSSSQ
DSENITAAALATGACIVGILCLPLILLLVYKQRQAASNRRAQELVRMDSNIQGIENPGFEAS
PPAQGIPEAKVRHPLSYVAQRQPSESGRHLLSEPSTPLSPPGPGDVFFPSLDPVPDSPNFEVI
```

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 190-216

FIGURE 85

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCT

TTCCCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCTGCCCTCC

TTTCCTGCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGG

GGTCTGTGGGTTGATCTGTGGCCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCCTCCCGA

CTCCGCTCCCGGACCAGCGGCCTGACCCTGGGGAAAGG**ATG**GTTCCCGAGGTGAGGGTCCTC

TCCTCCTTGCTGGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCC

AGACATGTTCTGCCTTTTCCATGGGAAGAGATACTCCCCCGGCGAGAGCTGGCACCCCTACT

TGGAGCCACAAGGCCTGATGTACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGT

TGTTACCGCCTCCACTGTCCGCCTGTCCACTGCCCCCAGCCTGTGACGGAGCCACAGCAATG

CTGTCCCAAGTGTGTGGAACCTCACACTCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCC

AGCACAACGGGACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTTCCCCTCC

CGCCTGCCCAACCAGTGTGTCCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCTCAC

AACCTGCCCCGAACCAGGCTGCCCAGCACCCCTCCCACTGCCAGACTCCTGCTGCCAAGCCT

GCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTG

AGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAGAGAGGCCCGGGCACCCCAGC

CCCCACTGGCCTCAGCGCCCCTCTGAGCTTCATCCCTCGCCACTTCAGACCCAAGGGAGCAG

GCAGCACAACTGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGGCGGG

AAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGCCCCTTGCCCTG

CATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGT

ACCCCTGCCGTCACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTTGCCCAGAGGACAAA

GCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTCCT

CGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGG

CCTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAAACTGAGGCTCAG

AGAGGTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGA

AAGTCAGGAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCAC

GAAGGTCACTGGAACGTCTTCCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGA

CAAAGTGACCAAGACATAACAAAGACC**TAA**CAGTTGCAGATATGAGCTGTATAATTGTTGTT

ATTATATATTAATAAATAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 86

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64902
><subunit 1 of 1, 451 aa, 1 stop
><MW: 49675, pI: 7.15, NX(S/T): 1
MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCT
CSEGAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIF
SAHELFPSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEED
SVQSLHGVRHPQDPCSSDAGRKRGPGTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKH
KKACVHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKC
CKICPEDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLV
KDEETEAQRGEVPGPRPHSQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSPDPG
AEGHGQSRQSDQDITKT


Signal peptide:
amino acids 1-25
```

86/270

FIGURE 87

CTAGCCTGCGCCAAGGGGTAGTGAGACCGCGCGGCAACAGCTTGCGGCTGCGGGGAGCTCCC

GTGGGCGCTCCGCTGGCTGTGCAGGCGGCC<u>ATG</u>GATTCCTTGCGGAAAATGCTGATCTCAGT

CGCAATGCTGGGCGCAGGGGCTGGCGTGGGCTACGCGCTCCTCGTTATCGTGACCCCGGGAG

AGCGGCGGAAGCAGGAAATGCTAAAGGAGATGCCACTGCAGGACCCAAGGAGCAGGGAGGAG

GCGGCCAGGACCCAGCAGCTATTGCTGGCCACTCTGCAGGAGGCAGCGACCACGCAGGAGAA

CGTGGCCTGGAGGAAGAACTGGATGGTTGGCGGCGAAGGCGGCGCCAGCGGGAGGTCACCG<u>T</u>

<u>GA</u>GACCGGACTTGCCTCCGTGGGCGCCGGACCTTGGCTTGGGCGCAGGAATCCGAGGCAGCC

TTTCTCCTTCGTGGGCCCAGCGGAGAGTCCGGACCGAGATACCATGCCAGGACTCTCCGGGG

TCCTGTGAGCTGCCGTCGGGTGAGCACGTTTCCCCCAAACCCTGGACTGACTGCTTTAAGGT

CCGCAAGGCGGGCCAGGGCCGAGACGCGAGTCGGATGTGGTGAACTGAAAGAACCAATAAAA

TCATGTTCCTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAA

FIGURE 88

MDSLRKMLISVAMLGAGAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAARTQQLLLA

TLQEAATTQENVAWRKNWMVGGEGGASGRSP

Signal peptide:

amino acids 1-18

FIGURE 89

CAGGAGAGAAGGCACCGCCCCCACCCCGCCTCCAAAGCTAACCCTCGGGCTTGAGGGGAAGA

GGCTGACTGTACGTTCCTTCTACTCTGGCACCACTCTCCAGGCTGCC**ATG**GGGCCCAGCACC

CCTCTCCTCATCTTGTTCCTTTTGTCATGGTCGGGACCCCTCCAAGGACAGCAGCACCACCT

TGTGGAGTACATGGAACGCCGACTAGCTGCTTTAGAGGAACGGCTGGCCCAGTGCCAGGACC

AGAGTAGTCGGCATGCTGCTGAGCTGCGGGACTTCAAGAACAAGATGCTGCCACTGCTGGAG

GTGGCAGAGAAGGAGCGGGAGGCACTCAGAACTGAGGCCGACACCATCTCCGGGAGAGTGGA

TCGTCTGGAGCGGGAGGTAGACTATCTGGAGACCCAGAACCCAGCTCTGCCCTGTGTAGAGT

TTGATGAGAAGGTGACTGGAGGCCCTGGGACCAAAGGCAAGGGAAGAAGGAATGAGAAGTAC

GATATGGTGACAGACTGTGGCTACACAATCTCTCAAGTGAGATCAATGAAGATTCTGAAGCG

ATTTGGTGGCCCAGCTGGTCTATGGACCAAGGATCCACTGGGGCAAACAGAGAAGATCTACG

TGTTAGATGGGACACAGAATGACACAGCCTTTGTCTTCCCAAGGCTGCGTGACTTCACCCTT

GCCATGGCTGCCCGGAAAGCTTCCCGAGTCCGGGTGCCCTTCCCCTGGGTAGGCACAGGGCA

GCTGGTATATGGTGGCTTTCTTTATTTTGCTCGGAGGCCTCCTGGAAGACCTGGTGGAGGTG

GTGAGATGGAGAACACTTTGCAGCTAATCAAATTCCACCTGGCAAACCGAACAGTGGTGGAC

AGCTCAGTATTCCCAGCAGAGGGGCTGATCCCCCCCTACGGCTTGACAGCAGACACCTACAT

CGACCTGGTAGCTGATGAGGAAGGTCTTTGGGCTGTCTATGCCACCCGGGAGGATGACAGGC

ACTTGTGTCTGGCCAAGTTAGATCCACAGACACTGGACACAGAGCAGCAGTGGGACACACCA

TGTCCCAGAGAGAATGCTGAGGCTGCCTTTGTCATCTGTGGGACCCTCTATGTCGTCTATAA

CACCCGTCCTGCCAGTCGGGCCCGCATCCAGTGCTCCTTTGATGCCAGCGGCACCCTGACCC

CTGAACGGGCAGCACTCCCTTATTTTCCCCGCAGATATGGTGCCCATGCCAGCCTCCGCTAT

AACCCCCGAGAACGCCAGCTCTATGCCTGGGATGATGGCTACCAGATTGTCTATAAGCTGGA

GATGAGGAAGAAAGAGGAGGAGGTT**TGA**GGAGCTAGCCTTGTTTTTTGCATCTTTCTCACTC

CCATACATTTATATTATATCCCCACTAAATTTCTTGTTCCTCATTCTTCAAATGTGGGCCAG

TTGTGGCTCAAATCCTCTATATTTTTAGCCAATGGCAATCAAATTCTTTCAGCTCCTTTGTT

TCATACGGAACTCCAGATCCTGAGTAATCCTTTTAGAGCCCGAAGAGTCAAAACCCTCAATG

TTCCCTCCTGCTCTCCTGCCCCATGTCAACAAATTTCAGGCTAAGGATGCCCCAGACCCAGG

GCTCTAACCTTGTATGCGGGCAGGCCCAGGGAGCAGGCAGCAGTGTTCTTCCCCTCAGAGTG

ACTTGGGGAGGGAGAAATAGGAGGAGACGTCCAGCTCTGTCCTCTCTTCCTCACTCCTCCCT

TCAGTGTCCTGAGGAACAGGACTTTCTCCACATTGTTTTGTATTGCAACATTTTGCATTAAA

AGGAAAATCCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAA

89/270

FIGURE 90

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64905
<subunit 1 of 1, 406 aa, 1 stop
<MW: 46038, pI: 6.50, NX(S/T): 2
MGPSTPLLILFLLSWSGPLQGQQHHLVEYMERRLAALEERLAQCQDQSSRHAAELRDFKNKM
LPLLEVAEKEREALRTEADTISGRVDRLEREVDYLETQNPALPCVEFDEKVTGGPGTKGKGR
RNEKYDMVTDCGYTISQVRSMKILKRFGGPAGLWTKDPLGQTEKIYVLDGTQNDTAFVFPRL
RDFTLAMAARKASRVRVPFPWVGTGQLVYGGFLYFARRPPGRPGGGGEMENTLQLIKFHLAN
RTVVDSSVFPAEGLIPPYGLTADTYIDLVADEEGLWAVYATREDDRHLCLAKLDPQTLDTEQ
QWDTPCPRENAEAAFVICGTLYVVYNTRPASRARIQCSFDASGTLTPERAALPYFPRRYGAH
ASLRYNPRERQLYAWDDGYQIVYKLEMRKKEEEV
```

Important features:
Signal peptide:
amino acids 1-21

N-glycosylation sites.
amino acids 177-180, 248-251

90/270

FIGURE 91

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAG
AGCCTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCAT
CCGTCACCTCTCCTGTCATCCGTTTCCATGCCGTGAGGTCCATTCACAGAACACATCC**ATGG**
CTCTCATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTT
GGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGGAGGACGCAGCATTCTCCTGTTTCCTGTC
TCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGGCCAGTTCTCTAGCGTGG
TCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAAGGCAGG
ACAAAACTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTAC
TGTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCA
TCTGGGAGCTACAGGTGTCAGCACTGGGCTCAGTTCCTCTCATTTCCATCACGGGATATGTT
GATAGAGACATCCAGCTACTCTGTCAGTCCTCGGGCTGGTTCCCCCGGCCCACAGCGAAGTG
GAAAGGTCCACAAGGACAGGATTTGTCCACAGACTCCAGGACAAACAGAGACATGCATGGCC
TGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAACGCCGGGAGCATATCCTGTTCCATG
CGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAGGAGATACCTTTTTCGA
GCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTATTTTTTG
GCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCCAGTGGAAAATCCAGGCGGAACTGGAC
TGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGAC
TCTGGATCCAGAGACCGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACTGTAACCCATA
GAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCT
TCTCAGAGTTTCCAAGCAGGGAAACATTACTGGGAGGTGGACGGAGGACACAATAAAAGGTG
GCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCCCG
ATCATGGGTACTGGGTCCTCAGACTGAATGGAGAACATTTGTATTTCACATTAAATCCCCGT
TTTATCAGCGTCTTCCCCAGGACCCCACCTACAAAAATAGGGGTCTTCCTGGACTATGAGTG
TGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGT
TTGAAGGCTTATTGAGGCCCTACATTGAGTATCCGTCCTATAATGAGCAAAATGGAACTCCC
ATAGTCATCTGCCCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGC
AATCCCAGAGACAAGCAACAGTGAGTCCTCCTCACAGGCAACCACGCCCTTCCTCCCCAGGG
GTGAAATG**TAG**GATGAATCACATCCCACATTCTTCTTTAGGGATATTAAGGTCTCTCTCCCA
GATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCAGATGAAGGGGGACTGGCCTGTCC
ACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGGCTGACATTACATTT
AGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAGAACCG
TCAGGAATTCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGC
TTAGATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTA
AAAAAA

FIGURE 92

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSS

VVHLYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQK

AIWELQVSALGSVPLISITGYVDRDIQLLCQSSGWFPRPTAKWKGPQGQDLSTDSRTNRDMH

GLFDVEISLTVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLF

FGIVGLKIFFSKFQWKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVT

HRKAPQEVPHSEKRFTRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLS

PDHGYWVLRLNGEHLYFTLNPRFISVFPRTPPTKIGVFLDYECGTISFFNINDQSLIYTLTC

RFEGLLRPYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLP

RGEM

Signal peptide:
amino acids 1-17

Transmembrane domain:
amino acids 239-255

FIGURE 93

```
GCGATGGTGCGCCCGGTGGCGGTGGCGGCGGTTGCGGAGGCTTCCTTGGTCGGATTGCA
ACGAGGAGAAGATGACTGACCAACCGACTGGCTGAATGAATGAATGGCGGAGCCGAGCGCGC
CATGAGGAGCCTGCCGAGCCTGGGCGGCCTCGCCCTGTTGTGCTGCGCCGCCGCCGCCGCCG
CCGTCGCCTCAGCCGCCTCGGCGGGGAATGTCACCGGTGGCGGCGGGGCCGCGGGGCAGGTG
GACGCGTCGCCGGGCCCCGGGTTGCGGGGCGAGCCCAGCCACCCCTTCCCTAGGGCGACGGC
TCCCACGGCCCAGGCCCCGAGGACCGGGCCCCCGCGCGCCACCGTCCACCGACCCCTGGCTG
CGACTTCTCCAGCCCAGTCCCCGGAGACCACCCCTCTTTGGGCGACTGCTGGACCCTCTTCC
ACCACCTTTCAGGCGCCGCTCGGCCCCTCGCCGACCACCCCTCCGGCGGCGGAACGCACTTC
GACCACCTCTCAGGCGCCGACCAGACCCGCGCCGACCACCCTTTCGACGACCACTGGCCCGG
CGCCGACCACCCCTGTAGCGACCACCGTACCGGCGCCCACGACTCCCCGGACCCCGACCCCC
GATCTCCCCAGCAGCAGCAACAGCAGCGTCCTCCCCACCCCACCTGCCACCGAGGCCCCCTC
TTCGCCTCCTCCAGAGTATGTATGTAACTGCTCTGTGGTTGGAAGCCTGAATGTGAATCGCT
GCAACCAGACCACAGGGCAGTGTGAGTGTCGGCCAGGTTATCAGGGGCTTCACTGTGAAACC
TGCAAAGAGGGCTTTTACCTAAATTACACTTCTGGGCTCTGTCAGCCATGTGACTGTAGTCC
ACATGGAGCTCTCAGCATACCGTGCAACAGGTAAGCAACAGAGGGTGGAACTGAÀGTTTATT
TTATTTTAGCAAGGGAAAAAAAAAAGGCTGCTACTCTCAAGGACCATACTGGTTTAAACAAAG
GAGGATGAGGGTCATAGATTTACAAAATATTTTATATACTTTTATTCTCTTACTTTATATGT
TATATTTAATGTCAGGATTTAAAAAACATCTAATTTACTGATTTAGTTCTTCAAAAGCACTAG
AGTCGCCAATTTTTCTCTGGGATAATTTCTGTAAATTTCATGGGAAAAAATTATTGAAGAAT
AAATCTGCTTTCTGGAAGGGCTTTCAGGCATGAAACCTGCTAGGAGGTTTAGAAATGTTCTT
ATGTTTATTAATATACCATTGGAGTTTGAGGAAATTTGTTGTTTGGTTTATTTTTCTCTCTA
ATCAAAATTCTACATTTGTTTCTTTGGACATCTAAAGCTTAACCTGGGGGTACCCTAATTTA
TTTAACTAGTGGTAAGTAGACTGGTTTTACTCTATTTACCAGTACATTTTTGAGACCAAAAG
TAGATTAAGCAGGAATTATCTTTAAACTATTATGTTATTTGGAGGTAATTTAATCTAGTGGA
ATAATGTACTGTTATCTAAGCATTTGCCTTGTACTGCACTGAAAGTAATTATTCTTTGACCT
TATGTGAGGCACTTGGCTTTTTGTGGACCCCAAGTCAAAAAACTGAAGAGACAGTATTAAAT
AATGAAAAAAATAATGACAGGTTATACTCAGTGTAACCTGGGTATAACCCAAGATCTGCTGC
CACTTACGAGCTGTGTTCCTTGGGCAAGTAATTTCCTTTCACTGAGCTTGTTTCTTCTCAAG
GTTGTTGTGAAGATTAAATGAGTTGATATATATAAAATGCCTAGCACATGTCACTCAATAAA
TTCTGGTTTGTTTTAATTTCAAAGGAATATTATGGACTGAAATGAGAGAACATGTTTTAAGA
ACTTTTAGCTCCTTGACAAAGAAGTGCTTTATACTTTAGCACTAAATATTTTAAATGCTTTA
TAAATGATATTATACTGTTATGGAATATTGTATCATATTGTAGTTTATTAAAAAATGTAGAAG
AGGCTGGGCGCGGTGGCTCACGCCTGTAATCCTAGCACTTTGGGAGGCCAAGGCGGGTGGAT
CACTTGAGGCCAGGAGTTCTAGATGAGCCTGGCCAGCACAGTGAAACCCCGTCTCTACTAAA
AATACAAACAAATTAGCTGGGCGTGGTGGCACACACCTGTAGTCCCAGCTACTCGGGAGGCT
GAGGCAGGAGAATCGGTTGAACCCGGGAGGTGGAGGTTGCAGTGAGCTGAGATCGCGCCACT
GCACTCCAGCCTGGTGAGAGAGGGAGACTCTGTCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

93/270

FIGURE 94

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64952

><subunit 1 of 1, 258 aa, 1 stop

><MW: 25716, pI: 8.13, NX(S/T): 5

MRSLPSLGGLALLCCAAAAAAVASAASAGNVTGGGGAAGQVDASPGPGLRGEPSHPFPRATA

PTAQAPRTGPPRATVHRPLAATSPAQSPETTPLWATAGPSSTTFQAPLGPSPTTPPAAERTS

TTSQAPTRPAPTTLSTTTGPAPTTPVATTVPAPTTPRTPTPDLPSSSNSSVLPTPPATEAPS

SPPPEYVCNCSVVGSLNVNRCNQTTGQCECRPGYQGLHCETCKEGFYLNYTSGLCQPCDCSP

HGALSIPCNR

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 30-33, 172-175, 195-198, 208-211, 235-238

EGF-like domain cysteine pattern signature.

amino acids 214-226.

FIGURE 95

TGCGGCGCAGTGTAGACCTGGGAGG**ATG**GGCGGCCTGCTGCTGGCTGCTTTTCTGGCTTTGG

TCTCGGTGCCCAGGGCCCAGGCCGTGTGGTTGGGAAGACTGGACCCTGAGCAGCTTCTTGGG

CCCTGGTACGTGCTTGCGGTGGCCTCCCGGGAAAAGGGCTTTGCCATGGAGAAGGACATGAA

GAACGTCGTGGGGGTGGTGGTGACCCTCACTCCAGAAAACAACCTGCGGACGCTGTCCTCTC

AGCACGGGCTGGGAGGGTGTGACCAGAGTGTCATGGACCTGATAAAGCGAAACTCCGGATGG

GTGTTTGAGAATCCCTCAATAGGCGTGCTGGAGCTCTGGGTGCTGGCCACCAACTTCAGAGA

CTATGCCATCATCTTCACTCAGCTGGAGTTCGGGGACGAGCCCTTCAACACCGTGGAGCTGT

ACAGTCTGACGGAGACAGCCAGCCAGGAGGCCATGGGGCTCTTCACCAAGTGGAGCAGGAGC

CTGGGCTTCCTGTCACAG**TAG**CAGGCCCAGCTGCAGAAGGACCTCACCTGTGCTCACAAGAT

CCTTCTGTGAGTGCTGCGTCCCCAGTAGGGATGGCGCCCACAGGGTCCTGTGACCTCGGCCA

GTGTCCACCCACCTCGCTCAGCGGCTCCCGGGGCCCAGCACCAGCTCAGAATAAAGCGATTC

CACAGCA

FIGURE 96

MGGLLLAAFLALVSVPRAQAVWLGRLDPEQLLGPWYVLAVASREKGFAMEKDMKNVVGVVVT

LTPENNLRTLSSQHGLGGCDQSVMDLIKRNSGWVFENPSIGVLELWVLATNFRDYAIIFTQL

EFGDEPFNTVELYSLTETASQEAMGLFTKWSRSLGFLSQ

Signal peptide:

amino acids 1-20

FIGURE 97

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCCAGAC<u>ATG</u>CTGCTGCTGCTGCTGCC

CCTGCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTT

CCGTGACGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCAT

GGCTGGATTTACCCTGGCCCAGTAGTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGA

CCAGGATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACC

GATTCCACCTCCTTGGGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGA

AGAAGTGATGCGGGGAGATACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAA

ACATCACCGGCTCTCTGTGAATGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAG

GCACCCTGGAGTCCGGCTGCCCCCAGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAG

GGGACACCCCCTATGATCTCCTGGATAGGGACCTCCGTGTCCCCCCTGGACCCCTCCACCAC

CCGCTCCTCGGTGCTCACCCTCATCCCACAGCCCCAGGACCATGGCACCAGCCTCACCTGTC

AGGTGACCTTCCCTGGGGCCAGCGTGACCACGAACAAGACCGTCCATCTCAACGTGTCCTAC

CCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGAGACGGCACAGTATCCACAGTCTTGGG

AAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCTGCGCCTGGTCTGTGCAGTTGATG

CAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGAGAGGCCTGACCCTGTGCCCC

TCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCACCTGAGGGATGCAGCTGA

ATTCACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCTGAACGTCTCCCTGC

AGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAGCCACAGCCCTG

GTCTTCCTGTCCTTCTGCGTCATCTTCGTTGTAGTGAGGTCCTGCAGGAAGAAATCGGCAAG

GCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTCAGCCT

CTCAGGGGCCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAGCT

TCTGCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGT

GAAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCC

ACAGA<u>TGA</u>GAAACTGCAGAGACTCACCCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGA

GAAGTCAGAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACT

ATGAATTATGTGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAAT

CCACACTGTGCCCTCCCTTTTATTTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 98

MLLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWF
REGANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKG
SIKWNYKHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSVPWACEQGTPPMISWIGTSVS
PLDPSTTRSSVLTLIPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLTMTVFQGDG
TVSTVLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWV
HLRDAAEFTCRAQNPLGSQQVYLNVSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRS
CRKKSARPAAGVGDTGIEDANAVRGSASQGPLTEPWAEDSPPDQPPPASARSSVGEGELQYA
SLSFQMVKPWDSRGQEATDTEYSEIKIHR

Signal peptide:
amino acids 1-15

Transmembrane domain:
amino acids 351-370

FIGURE 99

GACGCCCAGTGACCTGCCGAGGTCGGCAGCACAGAGCTCTGGAG<u>ATG</u>AAGACCCTGTTCCTG

GGTGTCACGCTCGGCCTGGCCGCTGCCCTGTCCTTCACCCTGGAGGAGGAGGATATCACAGG

GACCTGGTACGTGAAGGCCATGGTGGTCGATAAGGACTTTCCGGAGGACAGGAGGCCCAGGA

AGGTGTCCCCAGTGAAGGTGACAGCCCTGGGCGGTGGGAAGTTGGAAGCCACGTTCACCTTC

ATGAGGGAGGATCGGTGCATCCAGAAGAAAATCCTGATGCGGAAGACGGAGGAGCCTGGCAA

ATACAGCGCCTATGGGGGCAGGAAGCTCATGTACCTGCAGGAGCTGCCCAGGAGGGACCACT

ACATCTTTTACTGCAAAGACCAGCACCATGGGGGCCTGCTCCACATGGGAAAGCTTGTGGGT

AGGAATTCTGATACCAACCGGGAGGCCCTGGAAGAATTTAAGAAATTGGTGCAGCGCAAGGG

ACTCTCGGAGGAGGACATTTTCACGCCCCTGCAGACGGGAAGCTGCGTTCCCGAACAC<u>TAG</u>G

CAGCCCCGGGTCTGCACCTCCAGAGCCCACCCTACCACCAGACACAGAGCCCGGACCACCT

GGACCTACCCTCCAGCCATGACCCTTCCCTGCTCCCACCCACCTGACTCCAAATAAAGTCCT

TTTCCCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65404

<subunit 1 of 1, 170 aa, 1 stop

<MW: 19457, pI: 9.10, NX(S/T): 0

MKTLFLGVTLGLAAALSFTLEEEDITGTWYVKAMVVDKDFPEDRRPRKVSPVKVTALGGGKL

EATFTFMREDRCIQKKILMRKTEEPGKYSAYGGRKLMYLQELPRRDHYIFYCKDQHHGGLLH

MGKLVGRNSDTNREALEEFKKLVQRKGLSEEDIFTPLQTGSCVPEH

Important features:

Signal peptide:

amino acids 1-17

100/270

FIGURE 101

GTTCCGCAGATGCAGAGGTTGAGGTGGCTGCGGGACTGGAAGTCATCGGGCAGAGGTCTCAC

AGCAGCCAAGGAACCTGGGGCCCGCTCCTCCCCCCTCCAGGCC<u>ATG</u>AGGATTCTGCAGTTAA

TCCTGCTTGCTCTGGCAACAGGGCTTGTAGGGGGAGAGACCAGGATCATCAAGGGGTTCGAG

TGCAAGCCTCACTCCCAGCCCTGGCAGGCAGCCCTGTTCGAGAAGACGCGGCTACTCTGTGG

GGCGACGCTCATCGCCCCCAGATGGCTCCTGACAGCAGCCCACTGCCTCAAGCCCCGCTACA

TAGTTCACCTGGGGCAGCACAACCTCCAGAAGGAGGAGGGCTGTGAGCAGACCCGGACAGCC

ACTGAGTCCTTCCCCCACCCCGGCTTCAACAACAGCCTCCCCAACAAAGACCACCGCAATGA

CATCATGCTGGTGAAGATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCC

TCTCCTCACGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCAGCACGTCC

AGCCCCCAGTTACGCCTGCCTCACACCTTGCGATGCGCCAACATCACCATCATTGAGCACCA

GAAGTGTGAGAACGCCTACCCCGGCAACATCACAGACACCATGGTGTGTGCCAGCGTGCAGG

AAGGGGGCAAGGACTCCTGCCAGGGTGACTCCGGGGGCCCTCTGGTCTGTAACCAGTCTCTT

CAAGGCATTATCTCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGTCTACAC

GAAAGTCTGCAAATATGTGGACTGGATCCAGGAGACGATGAAGAACAAT<u>TAG</u>ACTGGACCCA

CCCACCACAGCCCATCACCCTCCATTTCCACTTGGTGTTTGGTTCCTGTTCACTCTGTTAAT

AAGAAACCCTAAGCCAAGACCCTCTACGAACATTCTTTGGGCCTCCTGGACTACAGGAGATG

CTGTCACTTAATAATCAACCTGGGGTTCGAAATCAGTGAGACCTGGATTCAAATTCTGCCTT

GAAATATTGTGACTCTGGGAATGACAACACCTGGTTTGTTCTCTGTTGTATCCCCAGCCCCA

AAGACAGCTCCTGGCCATATATCAAGGTTTCAATAAATATTTGCTAAATGAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 102

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65405
<subunit 1 of 1, 250 aa, 1 stop
<MW: 27466, pI: 8.87, NX(S/T): 4
MRILQLILLALATGLVGGETRIIKGFECKPHSQPWQAALFEKTRLLCGATLIAPRWLLTAAH
CLKPRYIVHLGQHNLQKEEGCEQTRTATESFPHPGFNNSLPNKDHRNDIMLVKMASPVSITW
AVRPLTLSSRCVTAGTSCLISGWGSTSSPQLRLPHTLRCANITIIEHQKCENAYPGNITDTM
VCASVQEGGKDSCQGDSGGPLVCNQSLQGIISWGQDPCAITRKPGVYTKVCKYVDWIQETMKNN
```

Important features:
Signal peptide:
amino acids 1-18

Serine proteases, trypsin family, histidine active site.
amino acids 58-63

N-glycosylation sites.
amino acids 99-102, 165-168, 181-184, 210-213

Glycosaminoglycan attachment site.
amino acids 145-148

Kringle domain proteins.
amino acids 197-209, 47-64

Serine proteases, trypsin family, histidine protein
amino acids 199-209, 47-63, 220-243

Apple domain proteins
amino acids 222-249, 189-222

102/270

FIGURE 103

GAGCAGTGTTCTGCTGGAGCCG<u>ATG</u>CCAAAAACCATGCATTTCTTATTCAGATTCATTGTTT
TCTTTTATCTGTGGGGCCTTTTTACTGCTCAGAGACAAAAGAAAGAGGAGAGCACCGAAGAA
GTGAAAATAGAAGTTTTGCATCGTCCAGAAAACTGCTCTAAGACAAGCAAGAAGGGAGACCT
ACTAAATGCCCATTATGACGGCTACCTGGCTAAAGACGGCTCGAAATTCTACTGCAGCCGGA
CACAAAATGAAGGCCACCCCAAATGGTTTGTTCTTGGTGTTGGGCAAGTCATAAAAGGCCTA
GACATTGCTATGACAGATATGTGCCCTGGAGAAAAGCGAAAAGTAGTTATACCCCCTTCATT
TGCATACGGAAAGGAAGGCTATGCAGAAGGCAAGATTCCACCGGATGCTACATTGATTTTTG
AGATTGAACTTTATGCTGTGACCAAAGGACCACGGAGCATTGAGACATTTAAACAAATAGAC
ATGGACAATGACAGGCAGCTCTCTAAAGCCGAGATAAACCTCTACTTGCAAAGGGAATTTGA
AAAAGATGAGAAGCCACGTGACAAGTCATATCAGGATGCAGTTTTAGAAGATATTTTTAAGA
AGAATGACCATGATGGTGATGGCTTCATTTCTCCCAAGGAATACAATGTATACCAACACGAT
GAACTA<u>TAG</u>CATATTTGTATTTCTACTTTTTTTTTTTAGCTATTTACTGTACTTTATGTATA
AAACAAAGTCACTTTTCTCCAAGTTGTATTTGCTATTTTTCCCCTATGAGAAGATATTTTGA
TCTCCCCAATACATTGATTTTGGTATAATAAATGTGAGGCTGTTTTGCAAACTTAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 104

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65406

<subunit 1 of 1, 222 aa, 1 stop

<MW: 25794, pI: 6.24, NX(S/T): 1

MPKTMHFLFRFIVFFYLWGLFTAQRQKKEESTEEVKIEVLHRPENCSKTSKKGDLLNAHYDG
YLAKDGSKFYCSRTQNEGHPKWFVLGVGQVIKGLDIAMTDMCPGEKRKVVIPPSFAYGKEGY
AEGKIPPDATLIFEIELYAVTKGPRSIETFKQIDMDNDRQLSKAEINLYLQREFEKDEKPRD
KSYQDAVLEDIFKKNDHDGDGFISPKEYNVYQHDEL

Important features:

Endoplasmic reticulum targeting sequence.
amino acids 219-222

N-glycosylation site.
amino acids 45-48

FKBP-type peptidyl-prolyl cis-trans isomerase
amino acids 87-223, 129-142

EF-hand calcium-binding domain proteins
amino acids 202-214, 195-214

FIGURE 105

CAGAA**ATG**CAGGGACCATTGCTTCTTCCAGGCCTCTGCTTTCTGCTGAGCCTCTTTGGAGCT

GTGACTCAGAAAACCAAAACTTCCTGTGCTAAGTGCCCCCCAAATGCTTCCTGTGTCAATAA

CACTCACTGCACCTGCAACCATGGATATACTTCTGGATCTGGGCAGAAACTATTCACATTCC

CCTTGGAGACATGTAACGCCAGGCATGGTGGCTCGCGCCTG**TAA**TCCCAGTTCTTTGGGAAG

CCAAGGCAGGTGGATCACCTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATAGTGAAAC

CCCGTGTCTACTAAAAATACAAAAATCAGCCGGGCGTGGTGGTGCATGCCTGCAATCCCAGT

TACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACTCAGGAGGCAGAAGTTGCAGTGAACCC

AGATCCTGCCATTGCACTCCAGCATGGATGACAGAGCAAGACTCCGTCTCAAAAAGAAAAGA

TAGTTTCTTGTTTCATTTCGCGACTGCCCTCTCAGTGTTTCCTGGGATCCCCTCCCAAATAA

AGTACTTATATTCTC

FIGURE 106

MQGPLLLPGLCFLLSLFGAVTQKTKTSCAKCPPNASCVNNTHCTCNHGYTSGSGQKLFTFPL
ETCNARHGGSRL

Signal peptide:
amino acids 1-18

FIGURE 107

CAAGCAGGTCATCCCCTTGGTGACCTTCAAAGAGAAGCAGAGAGGGCAGAGGTGGGGGGCAC

AGGGAAAGGGTGACCTCTGAGATTCCCCTTTTCCCCCAGACTTTGGAAGTGACCCACC**ATG**G

GGCTCAGCATCTTTTTGCTCCTGTGTGTGTTCTTGGGCTCAGCCAGGCAGCCACACCGAAGATT

TTCAATGGCACTGAGTGTGGGCGTAACTCACAGCCGTGGCAGGTGGGGCTGTTTGAGGGCAC

CAGCCTGCGCTGCGGGGGTGTCCTTATTGACCACAGGTGGGTCCTCACAGCGGCTCACTGCA

GCGGCAGCAGGTACTGGGTGCGCCTGGGGGAACACAGCCTCAGCCAGCTCGACTGGACCGAG

CAGATCCGGCACAGCGGCTTCTCTGTGACCCATCCCGGCTACCTGGGAGCCTCGACGAGCCA

CGAGCACGACCTCCGGCTGCTGCGGCTGCGCCTGCCCGTCCGCGTAACCAGCAGCGTTCAAC

CCCTGCCCCTGCCCAATGACTGTGCAACCGCTGGCACCGAGTGCCACGTCTCAGGCTGGGGC

ATCACCAACCACCCACGGAACCCATTCCCGGATCTGCTCCAGTGCCTCAACCTCTCCATCGT

CTCCCATGCCACCTGCCATGGTGTGTATCCCGGGAGAATCACGAGCAACATGGTGTGTGCAG

GCGGCGTCCCGGGGCAGGATGCCTGCCAGGGTGATTCTGGGGGCCCCCTGGTGTGTGGGGGA

GTCCTTCAAGGTCTGGTGTCCTGGGGGTCTGTGGGGCCCTGTGGACAAGATGGCATCCCTGG

AGTCTACACCTATATTTGCAAGTATGTGGACTGGATCCGGATGATCATGAGGAACAAC**TGA**C

CTGTTTCCTCCACCTCCACCCCCACCCCTTAACTTGGGTACCCCTCTGGCCCTCAGAGCACC

AATATCTCCTCCATCACTTCCCCTAGCTCCACTCTTGTTGGCCTGGGAACTTCTTGGAACTT

TAACTCCTGCCAGCCCTTCTAAGACCCACGAGCGGGGTGAGAGAAGTGTGCAATAGTCTGGA

ATAAATATAAATGAAGGAGGGGCAAAAAAAAAAAAAA

FIGURE 108

MGLSIFLLLCVLGLSQAATPKIFNGTECGRNSQPWQVGLFEGTSLRCGGVLIDHRWVLTAAH

CSGSRYWVRLGEHSLSQLDWTEQIRHSGFSVTHPGYLGASTSHEHDLRLLRLRLPVRVTSSV

QPLPLPNDCATAGTECHVSGWGITNHPRNPFPDLLQCLNLSIVSHATCHGVYPGRITSNMVC

AGGVPGQDACQGDSGGPLVCGGVLQGLVSWGSVGPCGQDGIPGVYTYICKYVDWIRMIMRNN

Signal peptide:
amino acids 1-17

FIGURE 109

GCGGCCACACGCAGCTAGCCGGAGCCCGGACCAGGCGCCTGTGCCTCCTCCTCGTCCCTCGC
CGCGTCCGCGAAGCCTGGAGCCGGCGGGAGCCCCGCGCTCGC**ATG**TCGGGCGAGCTCAGCA
ACAGGTTCCAAGGAGGGAAGGCGTTCGGCTTGCTCAAAGCCCGGCAGGAGAGGAGGCTGGCC
GAGATCAACCGGGAGTTTCTGTGTGACCAGAAGTACAGTGATGAAGAGAACCTTCCAGAAAA
GCTCACAGCCTTCAAAGAGAAGTACATGGAGTTTGACCTGAACAATGAAGGCGAGATTGACC
TGATGTCTTTAAAGAGGATGATGGAGAAGCTTGGTGTCCCCAAGACCCACCTGGAGATGAAG
AAGATGATCTCAGAGGTGACAGGAGGGGTCAGTGACACTATATCCTACCGAGACTTTGTGAA
CATGATGCTGGGGAAACGGTCGGCTGTCCTCAAGTTAGTCATGATGTTTGAAGGAAAAGCCA
ACGAGAGCAGCCCCAAGCCAGTTGGCCCCCCTCCAGAGAGAGACATTGCTAGCCTGCCC**TGA**
GGACCCCGCCTGGACTCCCCAGCCTTCCCACCCCATACCTCCCTCCCGATCTTGCTGCCCTT
CTTGACACACTGTGATCTCTCTCTCTCATTTGTTTGGTCATTGAGGGTTTGTTTGTGTTT
TCATCAATGTCTTTGTAAAGCACAAATTATCTGCCTTAAAGGGGCTCTGGGTCGGGGAATCC
TGAGCCTTGGGTCCCCTCCCTCTCTTCTTCCCTCCTTCCCCGCTCCCTGTGCAGAAGGGCTG
ATATCAAACCAAAAACTAGAGGGGGCAGGGCCAGGGCAGGGAGGCTTCCAGCCTGTGTTCCC
CTCACTTGGAGGAACCAGCACTCTCCATCCTTTCAGAAAGTCTCCAAGCCAAGTTCAGGCTC
ACTGACCTGGCTCTGACGAGGACCCCAGGCCACTCTGAGAAGACCTTGGAGTAGGGACAAGG
CTGCAGGGCCTCTTTCGGGTTTCCTTGGACAGTGCCATGGTTCCAGTGCTCTGGTGTCACCC
AGGACACAGCCACTCGGGGCCCCGCTGCCCCAGCTGATCCCCACTCATTCCACACCTCTTCT
CATCCTCAGTGATGTGAAGGTGGGAAGGAAAGGAGCTTGGCATTGGGAGCCCTTCAAGAAGG
TACCAGAAGGAACCCTCCAGTCCTGCTCTCTGGCCACACCTGTGCAGGCAGCTGAGAGGCAG
CGTGCAGCCCTACTGTCCCTTACTGGGGCAGCAGAGGGCTTCGGAGGCAGAAGTGAGGCCTG
GGGTTTGGGGGGAAAGGTCAGCTCAGTGCTGTTCCACCTTTTAGGGAGGATACTGAGGGGAC
CAGGATGGGAGAATGAGGAGTAAAATGCTCACGGCAAAGTCAGCAGCACTGGTAAGCCAAGA
CTGAGAAATACAAGGTTGCTTGTCTGACCCCAATCTGCTTGAAAAAAAAAAAAAAAAAAA

FIGURE 110

MSGELSNRFQGGKAFGLLKARQERRLAEINREFLCDQKYSDEENLPEKLTAFKEKYMEFDLN

NEGEIDLMSLKRMMEKLGVPKTHLEMKKMISEVTGGVSDTISYRDFVNMMLGKRSAVLKLVM

MFEGKANESSPKPVGPPPERDIASLP

FIGURE 111A

CGCGCTCCCCGCGCGCCTCCTCGGGCTCCACGCGTCTTGCCCCGCAGAGGCAGCCTCCTCCA
GGAGCGGGGCCCTGCACACC**ATG**GCCCCCGGGTGGGCAGGGGTCGGCGCCGCCGTGCGCGCC
CGCCTGGCGCTGGCCTTGGCGCTGGCGAGCGTCCTGAGTGGGCCTCCAGCCGTCGCCTGCCC
CACCAAGTGTACCTGCTCCGCTGCCAGCGTGGACTGCCACGGGCTGGGCCTCCGCGCGGTTC
CTCGGGGCATCCCCCGCAACGCTGAGCGCCTTGACCTGGACAGAAATAATATCACCAGGATC
ACCAAGATGGACTTCGCTGGGCTCAAGAACCTCCGAGTCTTGCATCTGGAAGACAACCAGGT
CAGCGTCATCGAGAGAGGCGCCTTCCAGGACCTGAAGCAGCTAGAGCGACTGCGCCTGAACA
AGAATAAGCTGCAAGTCCTTCCAGAATTGCTTTTCCAGAGCACGCCGAAGCTCACCAGACTA
GATTTGAGTGAAAACCAGATCCAGGGGATCCCGAGGAAGGCGTTCCGCGGCATCACCGATGT
GAAGAACCTGCAACTGGACAACAACCACATCAGCTGCATTGAAGATGGAGCCTTCCGAGCGC
TGCGCGATTTGGAGATCCTTACCCTCAACAACAACAACATCAGTCGCATCCTGGTCACCAGC
TTCAACCACATGCCGAAGATCCGAACTCTGCGCCTCCACTCCAACCACCTCTACTGCGACTG
CCACCTGGCCTGGCTCTCGGATTGGCTGCGACAGCGACGGACAGTTGGCCAGTTCACACTCT
GCATGGCTCCTGTGCATTTGAGGGGCTTCAACGTGGCGGATGTGCAGAAGAAGGAGTACGTG
TGCCCAGCCCCCCACTCGGAGCCCCCATCCTGCAATGCCAACTCCATCTCCTGCCCCTTCGCC
CTGCACGTGCAGCAATAACATCGTGGACTGTCGAGGAAAGGGCTTGATGGAGATTCCTGCCA
ACTTGCCGGAGGGCATCGTCGAAATACGCCTAGAACAGAACTCCATCAAAGCCATCCCTGCA
GGAGCCTTCACCCAGTACAAGAAACTGAAGCGAATAGACATCAGCAAGAATCAGATATCGGA
TATTGCTCCAGATGCCTTCCAGGGCCTGAAATCACTCACATCGCTGGTCCTGTATGGGAACA
AGATCACCGAGATTGCCAAGGGACTGTTTGATGGGCTGGTGTCCCTACAGCTGCTCCTCCTC
AATGCCAACAAGATCAACTGCCTGCGGGTGAACACGTTTCAGGACCTGCAGAACCTCAACTT
GCTCTCCCTGTATGACAACAAGCTGCAGACCATCAGCAAGGGGCTCTTCGCCCCTCTGCAGT
CCATCCAGACACTCCACTTAGCCCAAAACCCATTTGTGTGCGACTGCCACTTGAAGTGGCTG
GCCGACTACCTCCAGGACAACCCCATCGAGACAAGCGGGGCCCGCTGCAGCAGCCCGCGCCG
ACTCGCCAACAAGCGCATCAGCCAGATCAAGAGCAAGAAGTTCCGCTGCTCAGGCTCCGAGG
ATTACCGCAGCAGGTTCAGCAGCGAGTGCTTCATGGACCTCGTGTGCCCCGAGAAGTGTCGC
TGTGAGGGCACGATTGTGGACTGCTCCAACCAGAAGCTGGTCCGCATCCCAAGCCACCTCCC
TGAATATGTCACCGACCTGCGACTGAATGACAATGAGGTATCTGTTCTGGAGGCCACTGGCA
TCTTCAAGAAGTTGCCCAACCTGCGGAAAATAAATCTGAGTAACAATAAGATCAAGGAGGTG
CGAGAGGGAGCTTTCGATGGAGCAGCCAGCGTGCAGGAGCTGATGCTGACAGGGAACCAGCT
GGAGACCGTGCACGGGCGCGTGTTCCGTGGCCTCAGTGGCCTCAAAAACCTTGATGCTGAGGA
GTAACTTGATCAGCTGTGTGTGAGTAATGACACCTTTGCCGGCCTGAGTTCGGTGAGACTGCTG
TCCCTCTATGACAATCGGATCACCACCATCACCCCTGGGGCCTTCACCACGCTTGTCTCCCT
GTCCACCATAAACCTCCTGTCCAACCCCTTCAACTGCAACTGCCACCTGGCCTGGCTCGGCA
AGTGGTTGAGGAAGAGGCGGATCGTCAGTGGGAACCCTAGGTGCCAGAAGCCATTTTTCCTC
AAGGAGATTCCCATCCAGGATGTGGCCATCCAGGACTTCACCTGTGATGGCAACGAGGAGAG
TAGCTGCCAGCTGAGCCCGCGCTGCCCGGAGCAGTGCACCTGTATGGAGACAGTGGTGCGAT
GCAGCAACAAGGGGCTCCGCGCCCTCCCCAGAGGCATGCCCAAGGATGTGACCGAGCTGTAC
CTGGAAGGAAACCACCTAACAGCCGTGCCCAGAGAGCTGTCCGCCCTCCGACACCTGACGCT
TATTGACCTGAGCAACAACAGCATCAGCATGCTGACCAATTACACCTTCAGTAACATGTCTC
ACCTCTCCACTCTGATCCTGAGCTACAACCGGCTGAGGTGCATCCCCGTCCACGCCTTCAAC
GGGCTGCGGTCCCTGCGAGTGCTAACCCTCCATGGCAATGACATTTCCAGCGTTCCTGAAGG
CTCCTTCAACGACCTCACATCTCTTTCCCATCTGGCGCTGGGAACCAACCCACTCCACTGTG
ACTGCAGTCTTCGGTGGCTGTCGGAGTGGGTGAAGGCGGGGTACAAGGAGCCTGGCATCGCC
CGCTGCAGTAGCCCTGAGCCCATGGCTGACAGGCTCCTGCTCACCACCCCAACCCACCGCTT
CCAGTGCAAAGGGCCAGTGGACATCAACATTGTGGCCAAATGCAATGCCTGCCTCTCCAGCC
CGTGCAAGAATAACGGGACATGCACCCAGGACCCTGTGGAGCTGTACCGCTGTGCCTGCCCC

111/270

FIGURE 111B

```
TACAGCTACAAGGGCAAGGACTGCACTGTGCCCATCAACACCTGCATCCAGAACCCCTGTCA
GCATGGAGGCACCTGCCACCTGAGTGACAGCCACAAGGATGGGTTCAGCTGCTCCTGCCCTC
TGGGCTTTGAGGGGCAGCGGTGTGAGATCAACCCAGATGACTGTGAGGACAACGACTGCGAA
AACAATGCCACCTGCGTGGACGGGATCAACAACTACGTGTGTATCTGTCCGCCTAACTACAC
AGGTGAGCTATGCGACGAGGTGATTGACCACTGTGTGCCTGAGCTGAACCTCTGTCAGCATG
AGGCCAAGTGCATCCCCCTGGACAAAGGATTCAGCTGCGAGTGTGTCCCTGGCTACAGCGGG
AAGCTCTGTGAGACAGACAATGATGACTGTGTGGCCCACAAGTGCCGCCACGGGGCCCAGTG
CGTGGACACAATCAATGGCTACACATGCACCTGCCCCCAGGGCTTCAGTGGACCCTTCTGTG
AACACCCCCACCCATGGTCCTACTGCAGACCAGCCCATGCGACCAGTACGAGTGCCAGAAC
GGGGCCCAGTGCATCGTGGTGCAGCAGGAGCCCACCTGCCGCTGCCCACCAGGCTTCGCCGG
CCCCAGATGCGAGAAGCTCATCACTGTCAACTTCGTGGGCAAAGACTCCTACGTGGAACTGG
CCTCCGCCAAGGTCCGACCCCAGGCCAACATCTCCCTGCAGGTGGCCACTGACAAGGACAAC
GGCATCCTTCTCTACAAAGGAGACAATGACCCCCTGGCACTGGAGCTGTACCAGGGCCACGT
GCGGCTGGTCTATGACAGCCTGAGTTCCCCTCCAACCACAGTGTACAGTGTGGAGACAGTGA
ATGATGGGCAGTTTCACAGTGTGGAGCTGGTGACGCTAAACCAGACCCTGAACCTAGTAGTG
GACAAAGGAACTCCAAAGAGCCTGGGGAAGCTCCAGAAGCAGCCAGCAGTGGGCATCAACAG
CCCCCTCTACCTTGGAGGCATCCCCACCTCCACCGGCCTCTCCGCCTTGCGCCAGGGCACGG
ACCGGCCTCTAGGCGGCTTCCACGGATGCATCCATGAGGTGCGCATCAACAACGAGCTGCAG
GACTTCAAGGCCCTCCCACCACAGTCCCTGGGGGTGTCACCAGGCTGCAAGTCCTGCACCGT
GTGCAAGCACGGCCTGTGCCGCTCCGTGGAGAAGGACAGCGTGGTGTGCGAGTGCCGCCCAG
GCTGGACCGGCCCACTCTGCGACCAGGAGGCCCGGGACCCCTGCCTCGGCCACAGATGCCAC
CATGGAAAATGTGTGGCAACTGGGACCTCATACATGTGCAAGTGTGCCGAGGGCTATGGAGG
GGACTTGTGTGACAACAAGAATGACTCTGCCAATGCCTGCTCAGCCTTCAAGTGTCACCATG
GGCAGTGCCACATCTCAGACCAAGGGGAGCCCTACTGCCTGTGCCAGCCCGGCTTTAGCGGC
GAGCACTGCCAACAAGAGAATCCGTGCCTGGGACAAGTAGTCCGAGAGGTGATCCGCCGCCA
GAAAGGTTATGCATCATGTGCCACAGCCTCCAAGGTGCCCATCATGGAATGTCGTGGGGGCT
GTGGGCCCCAGTGCTGCCAGCCCACCCGCAGCAAGCGGCGGAAATACGTCTTCCAGTGCACG
GACGGCTCCTCGTTTGTAGAAGAGGTGGAGAGACACTTAGAGTGCGGCTGCCTCGCGTGTTC
CTAAGCCCCTGCCCGCCTGCCTGCCACCTCTCGGACTCCAGCTTGATGGAGTTGGGACAGCC
ATGTGGGACCCCCTGGTGATTCAGCATGAAGGAAATGAAGCTGGAGAGGAAGGTAAAGAAGA
AGAGAATATTAAGTATATTGTAAAATAAACAAAAAATAGAACTTAAAAAAAAAAAAAAAAAAA
AAAAAA
```

FIGURE 112

MAPGWAGVGAAVRARLALALALASVLSGPPAVACPTKCTCSAASVDCHGLGLRAVPRGIPRN

AERLDLDRNNITRITKMDFAGLKNLRVLHLEDNQVSVIERGAFQDLKQLERLRLNKNKLQVL

PELLFQSTPKLTRLDLSENQIQGIPRKAFRGITDVKNLQLDNNHISCIEDGAFRALRDLEIL

TLNNNNISRILVTSFNHMPKIRTLRLHSNHLYCDCHLAWLSDWLRQRRTVGQFTLCMAPVHL

RGFNVADVQKKEYVCPAPHSEPPSCNANSISCPSPCTCSNNIVDCRGKGLMEIPANLPEGIV

EIRLEQNSIKAIPAGAFTQYKKLKRIDISKNQISDIAPDAFQGLKSLTSLVLYGNKITEIAK

GLFDGLVSLQLLLLNANKINCLRVNTFQDLQNLNLLSLYDNKLQTISKGLFAPLQSIQTLHL

AQNPFVCDCHLKWLADYLQDNPIETSGARCSSPRRLANKRISQIKSKKFRCSGSEDYRSRFS

SECFMDLVCPEKCRCEGTIVDCSNQKLVRIPSHLPEYVTDLRLNDNEVSVLEATGIFKKLPN

LRKINLSNNKIKEVREGAFDGAASVQELMLTGNQLETVHGRVFRGLSGLKTLMLRSNLISCV

SNDTFAGLSSVRLLSLYDNRITTITPGAFTTLVSLSTINLLSNPFNCNCHLAWLGKWLRKRR

IVSGNPRCQKPFFLKEIPIQDVAIQDFTCDGNEESSCQLSPRCPEQCTCMETVVRCSNKGLR

ALPRGMPKDVTELYLEGNHLTAVPRELSALRHLTLIDLSNNSISMLTNYTFSNMSHLSTLIL

SYNRLRCIPVHAFNGLRSLRVLTLHGNDISSVPEGSFNDLTSLSHLALGTNPLHCDCSLRWL

SEWVKAGYKEPGIARCSSPEPMADRLLLTTPTHRFQCKGPVDINIVAKCNACLSSPCKNNGT

CTQDPVELYRCACPYSYKGKDCTVPINTCIQNPCQHGGTCHLSDSHKDGFSCSCPLGFEGQR

CEINPDDCEDNDCENNATCVDGINNYVCICPPNYTGELCDEVIDHCVPELNLCQHEAKCIPL

DKGFSCECVPGYSGKLCETDNDDCVAHKCRHGAQCVDTINGYTCTCPQGFSGPFCEHPPPMV

LLQTSPCDQYECQNGAQCIVVQQEPTCRCPPGFAGPRCEKLITVNFVGKDSYVELASAKVRP

QANISLQVATDKDNGILLYKGDNDPLALELYQGHVRLVYDSLSSPPTTVYSVETVNDGQFHS

VELVTLNQTLNLVVDKGTPKSLGKLQKQPAVGINSPLYLGGIPTSTGLSALRQGTDRPLGGF

HGCIHEVRINNELQDFKALPPQSLGVSFGCKSCTVCKHGLCRSVEKDSVVCECRPGWTGPLC

DQEARDPCLGHRCHHGKCVATGTSYMCKCAEGYGGDLCDNKNDSANACSAFKCHHGQCHISD

QGEPYCLCQPGFSGEHCQQENPCLGQVVREVIRRQKGYASCATASKVPIMECRGGCGPQCCQ

PTRSKRRKYVFQCTDGSSFVEEVERHLECGCLACS

Signal peptide:
amino acids 1-27

FIGURE 113

GGATGCAGGACGCTCCCCTGAGCTGCCTGTCACCGACTAGGTGGAGCAGTGTTTCTTCCGCA

GACTCAACTGAGAAGTCAGCCTCTGGGGCAGGCACCAGGAATCTGCCTTTTCAGTTCTGTCT

CCGGCAGGCTTTGAGG**ATG**AAGGCTGCGGGCATTCTGACCCTCATTGGCTGCCTGGTCACAG

GCGCCGAGTCCAAAATCTACACTCGTTGCAAACTGGCAAAAATATTCTCGAGGGCTGGCCTG

GACAATTACTGGGGCTTCAGCCTTGGAAACTGGATCTGCATGGCATATTATGAGAGCGGCTA

CAACACCACAGCCCCGACGGTCCTGGATGACGGCAGCATCGACTATGGCATCTTCCAGATCA

ACAGCTTCGCGTGGTGCAGACGCGGAAAGCTGAAGGAGAACAACCACTGCCATGTCGCCTGC

TCAGCCTTGATCACTGATGACCTCACAGATGCAATTATCTGTGCCAGGAAAATTGTTAAAGA

GACACAAGGAATGAACTATTGGCAAGGCTGGAAGAAACATTGTGAGGGCAGAGACCTGTCCG

AGTGGAAAAAAGGCTGTGAGGTTTCC**TAA**ACTGGAACTGGACCCAGGATGCTTTGCAGCAAC

GCCCTAGGATTTGCAGTGAATGTCCAAATGCCTGTGTCATCTTGTCCCGTTTCCTCCCAATA

TTCCTTCTCAAACTTGGAGAGGGAAAATTAAGCTATACTTTTAAGAAAATAAATATTTCCAT

TTAAATGTC

FIGURE 114

MKAAGILTLIGCLVTGAESKIYTRCKLAKIFSRAGLDNYWGFSLGNWICMAYYESGYNTTAP
TVLDDGSIDYGIFQINSFAWCRRGKLKENNHCHVACSALITDDLTDAIICARKIVKETQGMN
YWQGWKKHCEGRDLSEWKKGCEVS

Signal peptide:
amino acids 1-19

FIGURE 115

CAGGCCATTTGCATCCCACTGTCCTTGTGTTCGGAGCCAGGCCACACCGTCCTCAGCAGTGT
CATGTGTTAAAAACGCCAAGCTGAATATATC**ATG**CCCCTATTAAAACTTGTACATGGCTCCC
CATTGGTTTTTGGAGAAAAGTTCAAGCTTTTTACCTTGGTGTCTGCCTGTATCCCAGTGTTC
AGGCTGGCTAGACGGCGGAAGAAGATCCTATTTTACTGTCACTTCCCAGATCTGCTTCTCAC
CAAGAGAGATTCTTTTCTTAAACGACTATACAGGGCCCCAATTGACTGGATAGAGGAATACA
CCACAGGCATGGCAGACTGCATCTTAGTCAACAGCCAGTTCACAGCTGCTGTTTTTAAGGAA
ACATTCAAGTCCCTGTCTCACATAGACCCTGATGTCCTCTATCCATCTCTAAATGTCACCAG
CTTTGACTCAGTTGTTCCTGAAAAGCTGGATGACCTAGTCCCCAAGGGGAAAAAATTCCTGC
TGCTCTCCATCAACAGATACGAAAGGAAGAAAAATCTGACTTTGGCACTGGAAGCCCTAGTA
CAGCTGCGTGGAAGATTGACATCCCAAGATTGGGAGAGGGTTCATCTGATCGTGGCAGGTGG
TTATGACGAGAGAGTCCTGGAGAATGTGGAACATTATCAGGAATTGAAGAAAATGGTCCAAC
AGTCCGACCTTGGCCAGTATGTGACCTTCTTGAGGTCTTTCTCAGACAAACAGAAAATCTCC
CTCCTCCACAGCTGCACGTGTGTGCTTTACACACCAAGCAATGAGCACTTTGGCATTGTCCC
TCTGGAAGCCATGTACATGCAGTGCCCAGTCATTGCTGTTAATTCGGGTGGACCCTTGGAGT
CCATTGACCACAGTGTCACAGGGTTTCTGTGTGAGCCTGACCCGGTGCACTTCTCAGAAGCA
ATAGAAAAGTTCATCCGTGAACCTTCCTTAAAAGCCACCATGGGCCTGGCTGGAAGAGCCAG
AGTGAAGGAAAAATTTTCCCCTGAAGCATTTACAGAACAGCTCTACCGATATGTTACCAAAC
TGCTGGTA**TAA**TCAGATTGTTTTTAAGATCTCCATTAATGTCATTTTTATGGATTGTAGACC
CAGTTTTGAAACCAAAAAAGAAACCTAGAATCTAATGCAGAAGAGATCTTTTAAAAAATAAA
CTTGAGTCTTGAATGTGAGCCACTTTCCTATATACCACACCTCCCTGTCCACTTTTCAGAAA
AACCATGTCTTTTATGCTATAATCATTCCAAATTTTGCCAGTGTTAAGTTACAAATGTGGTG
TCATTCCATGTTCAGCAGAGTATTTTAATTATATTTTCTCGGGATTATTGCTCTTCTGTCTA
TAAATTTTGAATGATACTGTGCCTTAATTGGTTTTCATAGTTTAAGTGTGTATCATTATCAA
AGTTGATTAATTTGGCTTCATAGTATAATGAGAGCAGGGCTATTGTAGTTCCCAGATTCAAT
CCACCGAAGTGTTCACTGTCATCTGTTAGGGAATTTTTGTTTGTCCTGTCTTTGCCTGGATC
CATAGCGAGAGTGCTCTGTATTTTTTTTAAGATAATTTGTATTTTTGCACACTGAGATATAA
TAAAAGGTGTTTATCATAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MPLLKLVHGSPLVFGEKFKLFTLVSACIPVFRLARRRKKILFYCHFPDLLLTKRDSFLKRLY
RAPIDWIEEYTTGMADCILVNSQFTAAVFKETFKSLSHIDPDVLYPSLNVTSFDSVVPEKLD
DLVPKGKKFLLLSINRYERKKNLTLALEALVQLRGRLTSQDWERVHLIVAGGYDERVLENVE
HYQELKKMVQQSDLGQYVTFLRSFSDKQKISLLHSCTCVLYTPSNEHFGIVPLEAMYMQCPV
IAVNSGGPLESIDHSVTGFLCEPDPVHFSEAIEKFIREPSLKATMGLAGRARVKEKFSPEAF
TEQLYRYVTKLLV

Signal peptide:
amino acids 1-15

117/270

FIGURE 117

GACTACGCCGATCCGAGACGTGGCTCCCTGGGCGGCAGAACCATGTTGGACTTCGCGATCTT

CGCCGTTACCTTCTTGCTGGCGTTGGTGGGAGCCGTGCTCTACCTCTATCCGGCTTCCAGAC

AAGCTGCAGGAATTCCAGGGATTACTCCAACTGAAGAAAAAGATGGTAATCTTCCAGATATT

GTGAATAGTGGAAGTTTGCATGAGTTCCTGGTTAATTTGCATGAGAGATATGGGCCTGTGGT

CTCCTTCTGGTTTGGCAGGCGCCTCGTGGTTAGTTTGGGCACTGTTGATGTACTGAAGCAGC

ATATCAATCCCAATAAGACATCGGACCCTTTTGAAACCATGCTGAAGTCATTATTAAGGTAT

CAATCTGGTGGTGGCAGTGTGAGTGAAAACCACATGAGGAAAAAATTGTATGAAAATGGTGT

GACTGATTCTCTGAAGAGTAACTTTGCCCTCCTCCTAAAGCTTTCAGAAGAATTATTAGATA

AATGGCTCTCCTACCCAGAGACCCAGCACGTGCCCCTCAGCCAGCATATGCTTGGTTTTGCT

ATGAAGTCTGTTACACAGATGGTAATGGGTAGTACATTTGAAGATGATCAGGAGTCATTCG

CTTCCAGAAGAATCATGGCACAGTTTGGTCTGAGATTGGAAAAGGCTTTCTAGATGGGTCAC

TTGATAAAAACATGACTCGGAAAAAACAATATGAAGATGCCCTCATGCAACTGGAGTCTGTT

TTAAGGAACATCATAAAAGAACGAAAAGGAAGGAACTTCAGTCAACATATTTTCATTGACTC

CTTAGTACAAGGGAACCTTAATGACCAACAGATCCTAGAAGACAGTATGATATTTCTCTGG

CCAGTTGCATAATAACTGCAAAATTGTGTACCTGGGCAATCTGTTTTTTAACCACCTCTGAA

GAAGTTCAAAAAAAATTATATGAAGAGATAAACCAAGTTTTTGGAAATGGTCCTGTTACTCC

AGAGAAAATTGAGCAGCTCAGATATTGTCAGCATGTGCTTTGTGAAACTGTTCGAACTGCCA

AACTGACTCCAGTTTCTGCCCAGCTTCAAGATATTGAAGGAAAAATTGACCGATTTATTATT

CCTAGAGAGACCCTCGTCCTTTATGCCCTTGGTGTGGTACTTCAGGATCCTAATACTTGGCC

ATCTCCACACAAGTTTGATCCAGATCGGTTTGATGATGAATTAGTAATGAAAACTTTTTCCT

CACTTGGATTCTCAGGCACACAGGAGTGTCCAGAGTTGAGGTTTGCATATATGGTGACCACA

GTACTTCTTAGTGTATTGGTGAAGAGACTGCACCTACTTTCTGTGGAGGGACAGGTTATTGA

AACAAAGTATGAACTGGTAACATCATCAAGGGAAGAAGCTTGGATCACTGTCTCAAAGAGAT

ATTAAAATTTTATACATTTAAAATCATTGTTAAATTGATTGAGGAAACAACCATTTAAAAA

AAATCTATGTTGAATCCTTTTATAAACCAGTATCACTTTGTAATATAAACACCTATTTGTAC

TTAA

FIGURE 118

MLDFAIFAVTFLLALVGAVLYLYPASRQAAGIPGITPTEEKDGNLPDIVNSGSLHEFLVNLH
ERYGPVVSFWFGRRLVVSLGTVDVLKQHINPNKTSDPFETMLKSLLRYQSGGGSVSENHMRK
KLYENGVTDSLKSNFALLLKLSEELLDKWLSYPETQHVPLSQHMLGFAMKSVTQMVMGSTFE
DDQEVIRFQKNHGTVWSEIGKGFLDGSLDKNMTRKKQYEDALMQLESVLRNIIKERKGRNFS
QHIFIDSLVQGNLNDQQILEDSMIFSLASCIITAKLCTWAICFLTTSEEVQKKLYEEINQVF
GNGPVTPEKIEQLRYCQHVLCETVRTAKLTPVSAQLQDIEGKIDRFIIPRETLVLYALGVVL
QDPNTWPSPHKFDPDRFDDELVMKTFSSLGFSGTQECPELRFAYMVTTVLLSVLVKRLHLLS
VEGQVIETKYELVTSSREEAWITVSKRY

Signal peptide:
amino acids 1-18

Transmembrane domain:
amino acids 271-290

FIGURE 119

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGA

GACCGCCGCCCTTGTCCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTC

CTGACGCTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACAT

ACAGGCCTGCCTGCCTCTCACGTTCACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGG

TGGCCGCGCTCTCTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGA

GTCTCCATGTTCAACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGT

GGCCCTGTCCTTCTTCATATTCGAGCGTTGGGAGTGCACTACGTATTGGTACATTTTTGTCT

TCTGCAGTGCCCTTCCAGCTGTCACTGAAATGGCTTTATTCGTCACCGTCTTTGGGCTGAAA

AAGAAACCCTTCTGATTACCTTCATGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGG

CCGCTTCGTATTCCTGGAAGAAGGAAGGCATAGGCTTCGGTTTTCCCCTCGGAAACTGCTTC

TGCTGGAGGATATGTGTTGGAATAATTACGTCTTGAGTCTGGGATTATCCGCATTGTATTTA

GTGCTTTGTAATAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACA

ATTAAAAAAAAAAAA

FIGURE 120

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLG

LFAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVT

EMALFVTVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 121

TCCCGGACCCTGCCGCCCTGCCACT<u>ATG</u>TCCCGCCGCTCTATGCTGCTTGCCTGGGCTCTCC

CCAGCCTCCTTCGACTCGGAGCGGCTCAGGAGACAGAAGACCCGGCCTGCTGCAGCCCCATA

GTGCCCCGGAACGAGTGGAAGGCCCTGGCATCAGAGTGCGCCCAGCACCTGAGCCTGCCCTT

ACGCTATGTGGTGGTATCGCACACGGCGGGCAGCAGCTGCAACACCCCCGCCTCGTGCCAGC

AGCAGGCCCGGAATGTGCAGCACTACCACATGAAGACACTGGGCTGGTGCGACGTGGGCTAC

AACTTCCTGATTGGAGAAGACGGGCTCGTATACGAGGGCCGTGGCTGGAACTTCACGGGTGC

CCACTCAGGTCACTTATGGAACCCCATGTCCATTGGCATCAGCTTCATGGGCAACTACATGG

ATCGGGTGCCCACACCCCAGGCCATCCGGGCAGCCCAGGGTCTACTGGCCTGCGGTGTGGCT

CAGGGAGCCCTGAGGTCCAACTATGTGCTCAAAGGACACCGGGATGTGCAGCGTACACTCTC

TCCAGGCAACCAGCTCTACCACCTCATCCAGAATTGGCCACACTACCGCTCCCCC<u>TGA</u>GGCC

CTGCTGATCCGCACCCCATTCCTCCCCTCCCATGGCCAAAAACCCCACTGTCTCCTTCTCCA

ATAAAGATGTAGCTC

FIGURE 122

MSRRSMLLAWALPSLLRLGAAQETEDPACCSPIVPRNEWKALASΞCAQHLSLPLRYVVVSHT

AGSSCNTPASCQQQARNVQHYHMKTLGWCDVGYNFLIGEDGLVYΞGRGWNFTGAHSGHLWNP

MSIGISFMGNYMDRVPTPQAIRAAQGLLACGVAQGALRSNYVLKGHRDVQRTLSPGNQLYHL

IQNWPHYRSP

Signal peptide:
amino acids 1-20

FIGURE 123

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAG**ATG**CAACT

GACTCGCTGCTGCTTCGTGTTCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGG

ATGATGGTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCC

CGGGTGCCTCGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCT

CCTAGGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCCAACCGCC

CGAACCACAGCCCCCCACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTAC

TCCAACATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAA

TGGGACCTTCAGCGTCCACTTCCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCC

TCGTGCCCCCCAGTAAAGCTGTAGAGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAG

GCCTCCAAAATCTTCAACTGCCGGATGGAGTGGGAGAAGGTAGAACGGGGCCGCCGGACCTC

GCTTTGCACCCACGACCCAGCCAAGATCTGCTCCCGAGACCACGCTCAGAGCTCAGCCACCT

GGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTAT

CGGCTGGTCCAGAAGGTGTGCCCAGATTACAACTACCATAGTGATACCCCCTACTACCCATC

TGGG**TGA**CCCGGGGCAGGCCACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCCATGC

AGGAGACCATCTGGACACCGGGCAGGGAAGGGGTTGGGCCTCAGGCAGGGAGGGGGGTGGAG

ACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAGAGAAAGGGTCCCAAGTG

CTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGGAGGAGGAGTGGGCT

CTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGGTCCCCGAGGCC

TGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGGTT

CTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTCATCAGTGTGGACAGCCTG

TCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGGGC

CAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCTTGTCGTGTGCTGAGCATG

GCATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTC

CAGCCAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATT

GCTGATGGCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCCACACTAAGGCC

ACAGCCCATCCGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAG

CATCCATGTCCCGGAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTTCTCC

CGGATCTGGATGGCGCCGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGGCCGCAGA

GCATGTGCTGGATCTGTTCTGTGTGTCTGTCTGTGGGTGGGGGGAGGGGAGGGAAGTCTTGT

GAAACCGCTGATTGCTGACTTTTGTGTGAAGAATCGTGTTCTTGGAGCAGGAAATAAAGCTT

GCCCCGGGGCA

124/270

FIGURE 124

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66521

><subunit 1 of 1, 252 aa, 1 stop

><MW: 28127, pI: 8.91, NX(S/T): 5

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMAN

STLLGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVD

HGNGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGR

RTSLCTHDPAKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVCPDYNYHSDTPY

YPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 125

```
GTGAATGTGAGGGTTTGATGACTTTCAGATGTCTAGGAACCAGAGTGGGTGCAGGGGCCCCA
GGCAGGGCTGATTCTTGGGCGGAGGAGAGTAGGGTAAAGGGTTCTGCATGAGCTCCTTAAAG
GACAAAGGTAACAGAGCCAGCGAGAGAGCTCGAGGGGAGACTTTGACTTCAAGCCACAGAAT
TGGTGGAAGTGTGCGCGCCGCCGCCGCCGTCGCTCCTGCAGCGCTGTCGACCTAGCCGCTAG
CATCTTCCCGAGCACCGGGATCCCGGGGTAGGAGGCGACGCGGGCGAGCACCAGCGCCAGCC
GGCTGCGGCTGCCCACACGGCTCACCATGGGCTCCGGGCGCCGGGCGCTGTCCGCGGTGCCG
GCCGTGCTGCTGGTCCTCACGCTGCCGGGGCTGCCCGTCTGGGCACAGAACGACACGGAGCC
CATCGTGCTGGAGGGCAAGTGTCTGGTGGTGTGCGACTCGAACCCGGCCACGGACTCCAAGG
GCTCCTCTTCCTCCCCGCTGGGGATATCGGTCCGGGCGGCCAACTCCAAGGTCGCCTTCTCG
GCGGTGCGGAGCACCAACCACGAGCCATCCGAGATGAGCAACAAGACGCGCATCATTTACTT
CGATCAGATCCTGGTGAATGTGGGTAATTTTTTCACATTGGAGTCTGTCTTTGTAGCACCAA
GAAAAGGAATTTACAGTTTCAGTTTTCACGTGATTAAAGTCTACCAGAGCCAAACTATCCAG
GTTAACTTGATGTTAAATGGAAAACCAGTAATATCTGCCTTTGCGGGGGACAAAGATGTTAC
TCGTGAAGCTGCCACGAATGGTGTCCTGCTCTACCTAGATAAAGAGGATAAGGTTTACCTAA
AACTGGAGAAAGGTAATTTGGTTGGAGGCTGGCAGTATTCCACGTTTTCTGGCTTTCTGGTG
TTCCCCCTATAGGATTCAATTTCTCCATGATGTTCATCCAGGTGAGGGATGACCCACTCCTG
AGTTATTGGAAGATCATTTTTTCATCATTGGATTGATGTCTTTTATTGGTTTCTCATGGGTG
GATATGGATTCTAAGGATTCTAGCCTGTCTGAACCAATACAAAATTTCACAGATTATTTGTG
TGTGTCTGTTTCAGTATATATTTGGATTGGGACTCTAAGCAGATAATACCTATGCTTAAATGTA
ACAGTCAAAAGCTGTCTGCAAGACTTATTCTGAATTTCATTTCCTGGGATTACTGAATTAGT
TACAGATGTGGAATTTTATTTGTTTAGTTTTAAAAGACTGGCAACCAGGTCTAAGGATTAGA
AAACTCTAAAGTTCTGACTTCAATCAACGGTTAGTGTGATACTGCCAAAGAACTGTATACTG
TGTTAATATATTGATTATATTTGTTTTTATTCCTTTGGAATTAGTTTGTTTGGTTCTTGTAA
AAAACTTGGATTTTTTTTTTCAGTAACTGGTATTATGTTTTCTCTTAAAATAAGGTAATGAA
TGGCTTGCCCACAAATTTACCTTGACTACGATATCATCGACATGACTTCTCTCAAAAAAAAA
GAATGCTTCATAGTTGTATTTTAATTGTATATGTGAAAGAGTCATATTTTCCAAGTTATATT
TTCTAAGAAGAAGAATAGATCATAAATCTGACAAGGAAAAAGTTGCTTACCCAAAATCTAAG
TGCTCAATCCCTGAGCCTCAGCAAAACAGCTCCCCTCCGAGGGAAATCTTATACTTTATTGC
TCAACTTTAATTAAAATGATTGATAATAACCACTTTATTAAAAAACCTAAGGTTTTTTTTTTT
TCCGTAGACATGACCACTTTATTAACTGGTGGTGGGATGCTGTTGTTTCTAATTATACCTAT
TTTTCAAGGCTTCTGTTGTATTTGAAGTATCATCTGGTTTTGCCTTAACTCTTTAAATTGTA
TATATTTATCTGTTTAGCTAATATTAAATTCAAATATCCCATATCTAAATTTAGTGCAATAT
CTTGTCTTTTGTATAGGTCATATGAATTCATAAAATTATTTATGTCTGTTATAGAATAAAGA
TTAATATATGTTAAAAAAA
```

FIGURE 126

MGSGRRALSAVPAVLLVLTLPGLPVWAQNDTEPIVLEGKCLVVCDSNPATDSKGSSSSPLGI

SVRAANSKVAFSAVRSTNHEPSEMSNKTRIIYFDQILVNVGNFFTLESVFVAPRKGIYSFSF

HVIKVYQSQTIQVNLMLNGKPVISAFAGDKDVTREAATNGVLLYLDKEDKVYLKLEKGNLVG

GWQYSTFSGFLVFPL

Signal peptide:
amino acids 1-27

FIGURE 127

CGGTGGCC<u>ATG</u>ACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTC

GCCCTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGC

TTTCTTCTGGTTGGTGTCTCTACTGATTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTA

TTGACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTC

TATATCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTT

GAAGAGTATAAACCCAGGTGAGACAGCACCCTCTATGCGACTGCTGGCCTATGTTTCTGGCT

TGGGCTTTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGG

CCAGGCACAGTGGGCATTCATGGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGAC

GCTGGTCATTATCTTGCTGCATGTATTCTGGGGCATTGTATTTTTTGATGGCTGTGAGAAGA

AAAAGTGGGGCATCCTCCTTATCGTTCTCCTGACCCACCTGCTGGTGTCAGCCCAGACCTTC

ATAAGTTCTTATTATGGAATAAACCTGGCGTCAGCATTTATAATCCTGGTGCTCATGGGCAC

CTGGGCATTCTTA<u>G</u>CTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTCTGCCAAG

ACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGA<u>TAA</u>CCTCAGGGAACCAGCACTTCCCAA

ACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCTTTTTCTGAAAATCCCTTTTTCTG

GTGGAATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 128

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66658

><subunit 1 of 1, 257 aa, 1 stop

><MW: 28472, pI: 9.33, NX(S/T): 0

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDN

KDGPTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGF

GIMSGVFSFVNTLSDSLGPGTVGIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKW

GILLIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKN

FLLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 129

CGGCAACCAGCCGCCGCCACCACCGCTGCCACTGCCGCCCTGCCGGGGCCATGTTCGCTCTGGGCTTGCCCTTCT
TGGTGCTCTTGGTGGCCTCGGTCGAGAGCCATCTGGGGGTTCTGGGGCCCAAGAACGTCTCGCAGAAAGACGCCG
AGTTTGAGCGCACCTACGTGGACGAGGTCAACAGCGAGCTGGTCAACATCTACACCTTCAACCATACTGTGACCC
GCAACAGGACAGAGGGCGTGCGTGTGTCTGTGAACGTCCTGAACAAGCAGAAGGGGGCGCCGTTGCTGTTTGTGG
TCCGCCAGAAGGAGGCTGTGGTGTCCTTCCAGGTGCCCCTAATCCTGCGAGGGATGTTTCAGCGCAAGTACCTCT
ACCAAAAAGTGGAACGAACCCTGTGTCAGCCCCCCACCAAGAATGAGTCGGAGATTCAGTTCTTCTACGTGGATG
TGTCCACCCTGTCACCAGTCAACACCACATACCAGCTCCGGGTCAGCCGCATGGACGATTTTGTGCTCAGGACTG
GGGAGCAGTTCAGCTTCAATACCACAGCAGCACAGCCCCAGTACTTCAAGTATGAGTTCCCTGAAGGCGTGGACT
CGGTAATTGTCAAGGTGACCTCCAACAAGGCCTTCCCCTGCTCAGTCATCTCCATTCAGGATGTGCTGTGTCCTG
TCTATGACCTGGACAACAACGTAGCCTTCATCGGCATGTACCAGACGATGACCAAGAAGGCGGCCATCACCGTAC
AGCGCAAAGACTTCCCCAGCAACAGCTTTTATGTGGTGGTGGTGGTGAAGACCGAAGACCAAGCCTGCGGGGGCT
CCCTGCCTTTCTACCCCTTCGCAGAAGATGAACCGGTCGATCAAGGGCACCGCCAGAAAACCCTGTCAGTGCTGG
TGTCTCAAGCAGTCACGTCTGAGGCATACGTCAGTGGGATGCTCTTTTGCCTGGGTATATTTCTCTCCTTTTACC
TGCTGACCGTCCTCCTGGCCTGCTGGGAGAACTGGAGGCAGAAGAAGAAGACCCTGCTGGTGGCCATTGACCGAG
CCTGCCCAGAAAGCGGTCACCCTCGAGTCCTGGCTGATTCTTTTCCTGGCAGTTCCCCTTATGAGGGTTACAACT
ATGGCTCCTTTGAGAATGTTTCTGGATCTACCGATGGTCTGGTTGACAGCGCTGGCACTGGGGACCTCTCTTACG
GTTACCAGGGCCGCTCCTTTGAACCTGTAGGTACTCGGCCCCGAGTGGACTCCATGAGCTCTGTGGAGGAGGATG
ACTACGACACATTGACCGACATCGATTCCGACAAGAATGTCATTCGCACCAAGCAATACCTCTATGTGGCTGACC
TGGCACGGGAAGGACAAGCGTGTTCTGCGGAAAAAGTACCAGATCTACTTCTGGAACATTGCCACCATTGCTGTCT
TCTATGCCCTTCCTGTGGTGCAGCTGGTGATCACCTACCAGACGGTGGTGAATGTCACAGGGAATCAGGACATCT
GCTACTACAACTTCCTCTGCGCCCACCCACTGGGCAATCTCAGCCGCCTTCAACAACATCCTCAGCAACCTGGGGT
ACATCCTGCTGGGGCTGCTTTTCCTGCTCATCATCCTGCAACGGGAGATCAACCACAACCGGGCCCTGCTGCGCA
ATGACCTCTGTGCCCTGGAATGTGGGATCCCCAAACACTTTGGGCTTTTCTACGCCATGGGCACAGCCCTGATGA
TGGAGGGGCTGCTCAGTGCTTGCTATCATGTGTGCCCCAACTATACCAATTTCCAGTTTGACACATCGTTCATGT
ACATGATCGCCGGACTCTGCATGCTGAAGCTCTACCAGAAGCGGCACCCGGACATCAACGCCAGCGCCTACAGTG
CCTACGCCTGCCTGGCCATTGTCATCTTCTTCTCTGTGCTGGGCGTGGTCTTTGGCAAAGGGAACACGGCGGTTCT
GGATCGTCTTCTCCATCATTCACATCATCGCCACCCTGCTCCTCAGCACGCAGCTCTATTACATGGGCCGGTGGA
AACTGGACTCGGGGATCTTCCGCCGCATCCTCCACGTGCTCTACACAGACTGCATCCGGCAGTGCAGCGGGCCGC
TCTACGTGGACCGCATGGTGCTGCTGGTCATGGGCAACGTCATCAACTGGTCGCTGGCTGCCTATGGGCTTATCA
TGCGCCCCAATGATTTCGCTTCCTACTTGTTGGCCATTGGCATCTGCAACCTGCTCCTTTACTTCGCCTTCTACA
TCATCATGAAGCTCCGGAGTGGGGAGAGGATCAAGCTCATCCCCCTGCTCTGCATCGTTTGCACCTCCGTGGTCT
GGGGGCTTCGCGCTCTTCTTCTTCTTCCAGGGACTCAGCACCTGGCAGAAAACCCCTGCAGAGTCGAGGGAGCACA
ACCGGGACTGCATCCTCCTCGACTTCTTTGACGACCACGACATCTGGCACTTCCTCTCCTCCATCGCCATGTTCG
GGTCCTTCCTGGTGTTGCTGACACTGGATGACGACCTGGATACTGTGCAGCGGGACAAGATCTATGTCTTCTAGC
AGGAGCTGGGCCCTTCGCTTCACCTCAAGGGGCCCTGAGCTCCTTTGTGTCATAGACCGGTCACTCTGTCGTGCT
GTGGGGATGAGTCCCAGCACCGCTGCCCAGCACTGGATGGCAGCAGGACAGCCAGGTCTAGCTTAGGCTTGGCCT
GGGACAGCCATGGGGTGGCATGGAACCTTGCAGCTGCCCTCTGCCGAGGAGCAGGCCTGCTCCCCTGGAACCCCC
AGATGTTGGCCAAATTGCTGCTTTCTTCTCAGTGTTGGGGCCTTCCATGGGCCCCTGTCCTTTGGCTCTCCATTT
GTCCCTTTGCAAGAGGAAGGATGGAAGGGACACCCTCCCCATTTCATGCCTTGCATTTTGCCCGTCCTCCTCCCC
ACAATGCCCCAGCCTGGGACCTAAGGCCTCTTTTTCCTCCCATACTCCCACTCCAGGGCCTAGTCTGGGGCCTGA
ATCTCTGTCCTGTATCAGGGCCCCAGTTCTCTTTGGGCTGTCCCTGGCTGCCATCACTACGCCCATTCCAGTCAGCC
AGGATGGATGGGGGTATGAGATTTTGGGGGTTGGCCAGCTGGTGCCAGACTTTTGGTGCTAAGGCCTGCAAGGGG
CCTGGGGCAGTGCGTATTCTCTTCCCTCTGACCTGTGCTCAGGGCTGGCTCTTTAGCAATGCGCTCAGCCCAATT
TGAGAACCGCCTTCTGATTCAAGAGGCTGAATTCAGAGGTCACCTCTTCATCCCATCAGCTCCCAGACTGATGCC
AGCACCAGGACTGGAGGGAGAAGCGCCTCACCCCTTCCCTTCCTTCTTTCCAGGCCCTTAGTCTTGCCAAACCCC
AGCTGGTGGCCTTTCAGTGCCATTGACACTGCCCAAGAATGTCCAGGGGCAAAGGAGGGATGATACAGAGTTCAG
CCCGTTCTGCCTCCACAGCTGTGGGCACCCCAGTGCCTACCTTAGAAAGGGGCTTCAGGAAGGGATGTGCTGTTT
CCCTCTACGTGCCCAGTCCTAGCCTCGCTCTAGGACCCAGGGCTGGCTTCTAAGTTTCCGTCCAGTCTTCAGGCA
AGTTCTGTGTTAGTCATGCACACACATACCTATGAAACCTTGGAGTTTACAAAGAATTGCCCCAGCTCTGGGCAC
CCTGGCCACCCTGGTCCTTGGATCCCCTTCGTCCCACCTGGTCCACCCCAGATGCTGAGGATGGGGGAGCTCAGG
CGGGGGCCTCTGCTTTGGGGATGGGAATGTGTTTTTCTCCCAAACTTGTTTTTTATAGCTCTGCTTGAAGGGCTGGG
AGATGAGGTGGGTCTGGATCTTTTCTCAGAGCGTCTCCATGCTATGGTTGCATTTCCGTTTTCTATGAATGAATT
TGCATTCAATAAACAACCAGACTCAAAAAAAAAAAAAAA

136/270

FIGURE 130

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66659

><subunit 1 of 1, 832 aa, 1 stop

><MW: 94454, pI: 6.94, NX(S/T): 12

MFALGLPFLVLLVASVESHLGVLGPKNVSQKDAEFERTYVDEVNSELVNIYTFNHTVTRNRT
EGVRVSVNVLNKQKGAPLLFVVRQKEAVVSFQVPLILRGMFQRKYLYQKVERTLCQPPTKNE
SEIQFFYVDVSTLSPVNTTYQLRVSRMDDFVLRTGEQFSFNTTAAQPQYFKYEFPEGVDSVI
VKVTSNKAFPCSVISIQDVLCPVYDLDNNVAFIGMYQTMTKKAAITVQRKDFPSNSFYVVVV
VKTEDQACGGSLPFYPFAEDEPVDQGHRQKTLSVLVSQAVTSEAYVSGMLFCLGIFLSFYLL
TVLLACWENWRQKKKTLLVAIDRACPESGHPRVLADSFPGSSPYEGYNYGSFENVSGSTDGL
VDSAGTGDLSYGYQGRSFEPVGTRPRVDSMSSVEEDDYDTLTDIDSDKNVIRTKQYLYVADL
ARKDKRVLRKKYQIYFWNIATIAVFYALPVVQLVITYQTVVNVTGNQDICYYNFLCAHPLGN
LSAFNNILSNLGYILLGLLFLLIILQREINHNRALLRNDLCALECGIPKHFGLFYAMGTALM
MEGLLSACYHVCPNYTNFQFDTSFMYMIAGLCMLKLYQKRHPDINASAYSAYACLAIVIFFS
VLGVVFGKGNTAFWIVFSIIHIIATLLLSTQLYYMGRWKLDSGIFRRILHVLYTDCIRQCSG
PLYVDRMVLLVMGNVINWSLAAYGLIMRPNDFASYLLAIGICNLLLYFAFYIIMKLRSGERI
KLIPLLCIVCTSVVWGFALFFFFQGLSTWQKTPAESREHNRDCILLDFFDDHDIWHFLSSIA
MFGSFLVLLTLDDDLDTVQRDKIYVF

Important features of the protein:

Signal peptide:
amino acids 1-18

Transmembrane domains:
amino acids 292-317, 451-470, 501-520, 607-627, 751-770

Leucine zipper pattern.
amino acids 497-518

N-glycosylation sites.
amino acids 27-30, 54-57, 60-63, 123-126, 141-144, 165-168, 364-367, 476-479, 496-499, 572-575, 603-606, 699-702

131/278

FIGURE 131

```
GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGC
TCTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTG
ACCATGGTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGC
CCAGCCTGCAGAGCTGTCTGTGGAAGTTCCAGAAAACTATGGTGGAAATTTCCCTTTATACC
TGACCAAGTTGCCGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTCAGGGGACTCA
GGCAAGGCAACTGAGGGCCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAG
GGCCCTGGACCGAGAGGAGCAGGCAGAGTACCAGCTACAGGTCACCCTGGAGATGCAGGATG
GACATGTCTTGTGGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTG
CCCCATTTCTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCC
CTTCCTCTTCCTTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACTCGGATCTTCGAT
TCCACATCCTGAGCCAGGCTCCAGCCCAGCCTTCCCCAGACATGTTCCAGCTGGAGCCTCGG
CTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGCACCAGCCTTGACCACGCCCTGGAGAGGAC
CTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACCAGGCCACTG
CCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGTCCCTAGAGCCTATCCACCTGGCA
GAGAATCTCAAAGTCCTATACCCGCACCACATGGCCCAGGTACACTGGAGTGGGGGTGATGT
GCACTATCACCTGGAGAGCCATCCCCCGGGACCCTTTGAAGTGAATGCAGAGGGAAACCTCT
ACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCCAGGTGCGGGCTCAG
AATTCCCATGGCGAGGACTATGCGGCCCCTCTGGAGCTGCACGTGCTGGTGATGGATGAGAA
TGACAACGTGCCTATCTGCCCTCCCCGTGACCCCACAGTCAGCATCCCTGAGCTCAGTCCAC
CAGGTACTGAAGTGACTAGACTGTCAGCAGAGGATGCAGATGCCCCCGGCTCCCCCAATTCC
CACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTCCA
GGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACA
TCCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGT
GAAGTCGAAGTCGCAGTCACAGATATCAATGATCACGCCCCTGAGTTCATCACTTCCCAGAT
TGGGCCTATAAGCCTCCCTGAGGATGTGGAGCCCGGGACTCTGGTGGCCATGCTAACAGCCA
TTGATGCTGACCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACA
GAAGGGACTTTTGGCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAA
GAACCTCAGTTATGAGGCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGA
AGCTGGTGGGGCCAGGCCCAGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGA
GTGATGCCACCCCCCAAGTTGGACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCC
AGCCGGCTCTTTCCTGCTGACCATCCAGCCCTCCGACCCCATCAGCCGAACCCTCAGGTTCT
CCCTAGTCAATGACTCAGAGGGCTGGCTCTGCATTGAGAAATTCTCCGGGGAGGTGCACACC
GCCCAGTCCCTGCAGGGCGCCCAGCCTGGGGACACCTACACGGTGCTTGTGGAGGCCCAGGA
TACAGCCCTGACTCTTGCCCCTGTGCCCTCCCAATACCTCTGCACACCCCGCCAAGACCATG
GCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCTGGCCAGTGGGCACGGTCCCTACAGC
TTCACCCTTGGTCCCAACCCCACGGTGCAACGGGATTGGCGCCTCCAGACTCTCAATGGTTC
CCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCACGTGAACACATAATCCCCGTGG
TGGTCAGCCACAATGCCCAGATGTGGCAGCTCCTGGTTCGAGTGATCGTGTGTCGCTGCAAC
GTGGAGGGGCAGTGCATGCGCAAGGTGGGCCGCATGAAGGGCATGCCCACGAAGCTGTCGGC
AGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAATCTTCCTCATCCTCATTTTCACCC
ACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCCTGAAGGCG
ACTGTCTGAATGGCCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAGTCCC
CTGGGAGAGAGCCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCCTCCAT
CTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCAAC
TTTATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTTTGCCCAATAATAAAGCCCCA
GAGAACTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG
```

132/270

FIGURE 132

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSG

KATEGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVP

HFSQAIYRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRL

GALALSPKGSTSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLEPIHLAE

NLKVLYPHHMAQVHWSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQN

SHGEDYAAPLELHVLVMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSH

VVYQLLSPEPEDGVEGRAFQVDPTSGSVTLGVLPLRAGQNILLLVLAMDLAGAEGGFSSTCE

VEVAVTDINDHAPEFITSQIGPISLPEDVEPGTLVAMLTAIDADLEPAFRLMDFAIERGDTE

GTFGLDWEPDSGHVRLRLCKNLSYEAAPSHEVVVVVQSVAKLVGPGPGPGATATVTVLVERV

MPPPKLDQESYEASVPISAPAGSFLLTIQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTA

QSLQGAQPGDTYTVLVEAQDTALTLAPVPSQYLCTPRQDHGLIVSGPSKDPDLASGHGPYSF

TLGPNPTVQRDWRLQTLNGSHAYLTLALHWVEPREHIIPVVVSHNAQMWQLLVRVIVCRCNV

EGQCMRKVGRMKGMPTKLSAVGILVGTLVAIGIFLILIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:
amino acids 1-18

Transmembrane domain:
amino acids 762-784

FIGURE 133

CCGGGGAC<u>ATG</u>AGGTGGATACTGTTCATTGGGGCCCTTATTGGGTCCAGCATCTGTGGCCAA
GAAAAATTTTTTGGGGACCAAGTTTTGAGGATTAATGTCAGAAATGGAGACGAGATCAGCAA
ATTGAGTCAACTAGTGAATTCAAACAACTTGAAGCTCAATTTCTGGAAATCTCCCTCCTCCT
TCAATCGGCCTGTGGATGTCCTGGTCCCATCTGTCAGTCTGCAGGCATTTAAATCCTTCCTG
AGATCCCAGGGCTTAGAGTACGCAGTGACAATTGAGGACCTGCAGGCCCTTTTAGACAATGA
AGATGATGAAATGCAACACAATGAAGGGCAAGAACGGAGCAGTAATAACTTCAACTACGGGG
CTTACCATTCCCTGGAAGCTATTTACCACGAGATGGACAACATTGCCGCAGACTTTCCTGAC
CTGGCGAGGAGGGTGAAGATTGGACATTCGTTTGAAAACCGGCCGATGTATGTACTGAAGTT
CAGCACTGGGAAAGGCGTGAGGCGGCCGGCCGTTTGGCTGAATGCAGGCATCCATTCCCGAG
AGTGGATCTCCCAGGCCACTGCAATCTGGACGGCAAGGAAGATTGTATCTGATTACCAGAGG
GATCCAGCTATCACCTCCATCTTGGAGAAAATGGATATTTTCTTGTTGCCTGTGGCCAATCC
TGATGGATATGTGTATACTCAAACTCAAAACCGATTATGGAGGAAGACGCGGTCCCGAAATC
CTGGAAGCTCCTGCATTGGTGCTGACCCAAATAGAAACTGGAACGCTAGTTTTGCAGGAAAG
GGAGCCAGCGACAACCCTTGCTCCGAAGTGTACCATGGACCCCACGCCAATTCGGAAGTGGA
GGTGAAATCAGTGGTAGATTTCATCCAAAAACATGGGAATTTCAAGGGCTTCATCGACCTGC
ACAGCTACTCGCAGCTGCTGATGTATCCATATGGGTACTCAGTCAAAAAGGCCCCAGATGCC
GAGGAACTCGACAAGGTGGCGAGGCTTGCGGCCAAAGCTCTGGCTTCTGTGTCGGGCACTGA
GTACCAAGTGGGTCCCACCTGCACCACTGTCTATCCAGCTAGCGGGAGCAGCATCGACTGGG
CGTATGACAACGGCATCAAATTTGCATTCACATTTGAGTTGAGAGATACCGGGACCTATGGC
TTCCTCCTGCCAGCTAACCAGATCATCCCCACTGCAGAGGAGACGTGGCTGGGGCTGAAGAC
CATCATGGAGCATGTGCGGGACAACCTCTAC<u>TAG</u>GCGATGGCTCTGCTCTGTCTACATTTAT
TTGTACCCACACGTGCACGCACTGAGGCCATTGTTAAAGGAGCTCTTTCCTACCTGTGTGAG
TCAGAGCCCTCTGGGTTTGTGGAGCACACAGGCCTGCCCCTCTCCAGCCAGCTCCCTGGAGT
CGTGTGTCCTGGCGGTGTCCCTGCAAGAACTGGTTCTGCCAGCCTGCTCAATTTTGGTCCTG
CTGTTTTTGATGAGCCTTTTGTCTGTTTCTCCTTCCACCCTGCTGGCTGGGCGGCTGCACTC
AGCATCACCCCTTCCTGGGTGGCATGTCTCTCTCTACCTCATTTTTAGAACCAAAGAACATC
TGAGATGATTCTCTACCCTCATCCACATCTAGCCAAGCCAGTGACCTTGCTCTGGTGGCACT
GTGGGAGACACCACTTGTCTTTAGGTGGGTCTCAAAGATGATGTAGAATTTCCTTTAATTTC
TCGCAGTCTTCCTGGAAAATATTTTCCTTTGAGCAGCAAATCTTGTAGGGATATCAGTGAAG
GTCTCTCCCTCCCTCCTCTCCTGTTTTTTTTTTTTTTGAGACAGAGTTTTGCTCTTGTTGCC
CAGGCTGGAGTGTGATGGCTCGATCTTGGCTCACCACAACCTCTGCCTCCTGGGTTCAAGCA
ATTCTCCTGCCTCAGCCTCTTGAGTAGCTTGGTTTATAGGCGCATGCCACCATGCCTGGCTA
ATTTTGTGTTTTTAGTAGAGACAGGGTTTCTCCATGTTGGTCAGGCTGGTCTCAAACTCCCA
ACCTCAGGTGATCTGCCCTCCTTGGCCTCCCAGAGTGCTGGGATTACAGGTGTGAGCCACTG
TGCCGGGCCCGTCCCCTCCTTTTTTAGGCCTGAATACAAAGTAGAAGATCACTTTCCTTCAC
TGTGCTGAGAATTTCTAGATACTACAGTTCTTACTCCTCTCTTCCCTTTGTTATTCAGTGTG
ACCAGGATGGCGGGAGGGGATCTGTGTCACTGTAGGTACTGTGCCCAGGAAGGCTGGGTGAA
GTGACCATCTAAATTGCAGGATGGTGAAATTATCCCCATCTGTCCTAATGGGCTTACCTCCT
CTTTGCCTTTTGAACTCACTTCAAAGATCTAGGCCTCATCTTACAGGTCCTAAATCACTCAT
CTGGCCTGGATAATCTCACTGCCCTGGCACATTCCCATTTGTGCTGTGGTGTATCCTGTGTT
TCCTTGTCCTGGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTTGTGTGTGTGTG
TCTGTCTATTTTGTATCCTGGACCACAAGTTCCTAAGTAGAGCAAGAATTCATCAACCAGCT
GCCTCTTGTTTCATTTCACCTCAGCACGTACCATCTGTCCTTTTGTTGTTGTTGTTTTGTTT
TTGTTTTTTTGCTTTTACCAAACATGTCTGTAAATCTTAACCTCCTGCCTAGGATTTGTACA
GCATCTGGTGTGTGCTTATAAGCCAATAAATATTCAATGTGAAAAAAAAAAAAAAAAAA

FIGURE 134

MRWILFIGALIGSSICGQEKFFGDQVLRINVRNGDEISKLSQLVNSNNLKLNFWKSPSSFNR

PVDVLVPSVSLQAFKSFLRSQGLEYAVTIEDLQALLDNEDDEMQHNEGQERSSNNFNYGAYH

SLEAIYHEMDNIAADFPDLARRVKIGHSFENRPMYVLKFSTGKGVRRPAVWLNAGIHSREWI

SQATAIWTARKIVSDYQRDPAITSILEKMDIFLLPVANPDGYVYTQTQNRLWRKTRSRNPGS

SCIGADPNRNWNASFAGKGASDNPCSEVYHGPHANSEVEVKSVVDFIQKHGNFKGFIDLHSY

SQLLMYPYGYSVKKAPDAEELDKVARLAAKALASVSGTEYQVGFTCTTVYPASGSSIDWAYD

NGIKFAFTFELRDTGTYGFLLPANQIIPTAEETWLGLKTIMEHVRDNLY

Signal peptide:
amino acids 1-16

FIGURE 135

CAACCATGCAAGGACAGGGCAGGAGAAGAGGAACCTGCAAAGACATATTTTGTTCCAAA**ATG**

GCATCTTACCTTTATGGAGTACTCTTTGCTGTTGGCCTCTGTGCTCCAATCTACTGTGTGTC

CCCGGCCAATGCCCCCAGTGCATACCCCCGCCCTTCCTCCACAAAGAGCACCCCTGCCTCAC

AGGTGTATTCCCTCAACACCGACTTTGCCTTCCGCCTATACCGCAGGCTGGTTTTGGAGACC

CCGAGTCAGAACATCTTCTTCTCCCCTGTGAGTGTCTCCACTTCCCTGGCCATGCTCTCCCT

TGGGGCCCACTCAGTCACCAAGACCCAGATTCTCCAGGGCCTGGGCTTCAACCTCACACACA

CACCAGAGTCTGCCATCCACCAGGGCTTCCAGCACCTGGTTCACTCACTGACTGTTCCCAGC

AAAGACCTGACCTTGAAGATGGGAAGTGCCCTCTTCGTCAAGAAGGAGCTGCAGCTGCAGGC

AAATTTCTTGGGCAATGTCAAGAGGCTGTATGAAGCAGAAGTCTTTTCTACAGATTTCTCCA

ACCCCTCCATTGCCCAGGCGAGGATCAACAGCCATGTGAAAAAGAAGACCCAAGGGAAGGTT

GTAGACATAATCCAAGGCCTTGACCTTCTGACGGCCATGGTTCTGGTGAATCACATTTTCTT

TAAAGCCAAGTGGGAGAAGCCCTTTCACCTTGAATATACAAGAAAGAACTTCCCATTCCTGG

TGGGCGAGCAGGTCACTGTGCAAGTCCCCATGATGCACCAGAAAGAGCAGTTCGCTTTTGGG

GTGGATACAGAGCTGAACTGCTTTGTGCTGCAGATGGATTACAAGGGAGATGCCGTGGCCTT

CTTTGTCCTCCCTAGCAAGGGCAAGATGAGGCAACTGGAACAGGCCTTGTCAGCCAGAACAC

TGATAAAGTGGAGCCACTCACTCCAGAAAAGGTGGATAGAGGTGTTCATCCCCAGATTTTCC

ATTTCTGCCTCCTACAATCTGGAAACCATCCTCCCGAAGATGGGCATCCAAAATGCCTTTGA

CAAAAATGCTGATTTTTCTGGAATTGCAAAGAGAGACTCCCTGCAGGTTTCTAAAGCAACCC

ACAAGGCTGTGCTGGATGTCAGTGAAGAGGGCACTGAGGCCACAGCAGCTACCACCACCAAG

TTCATAGTCCGATCGAAGGATGGTCCCTCTTACTTCACTGTCTCCTTCAATAGGACCTTCCT

GATGATGATTACAAATAAAGCCACAGACGGTATTCTCTTTCTAGGGAAAGTGGAAAATCCCA

CTAAATCC**TAG**GTGGGAAATGGCCTGTTAACTGATGGCACATTGCTAATGCACAAGAAATAA

CAAACCACATCCCTCTTTCTGTTCTGAGGGTGCATTTGACCCCAGTGGAGCTGGATTCGCTG

GCAGGGATGCCACTTCCAAGGCTCAATCACCAAACCATCAACAGGGACCCCAGTCACAAGCC

AACACCCATTAACCCCAGTCAGTGCCCTTTTCCACAAATTCTCCCAGGTAACTAGCTTCATG

GGATGTTGCTGGGTTACCATATTTCCATTCCTTGGGGCTCCCAGGAATGGAAATACGCCAAC

CCAGGTTAGGCACCTCTATTGCAGAATTACAATAACACATTCAATAAAACTAAAATATGAAT

TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAA

136/270

FIGURE 136

MASYLYGVLFAVGLCAPIYCVSPANAPSAYPRPSSTKSTPASQVYSLNTDFAFRLYRRLVLE

TPSQNIFFSPVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHSLTVP

SKDLTLKMGSALFVKKELQLQANFLGNVKRLYEAEVFSTDFSNPSIAQARINSHVKKKTQGK

VVDIIQGLDLLTAMVLVNHIFFKAKWEKPFHLEYTRKNFPPFLVGEQVTVQVPMMHQKEQFAF

GVDTELNCFVLQMDYKGDAVAFFVLPSKGKMRQLEQALSARTLIKWSHSLQKRWIEVFIPRF

SISASYNLETILPKMGIQNAFDKNADFSGIAKRDSLQVSKATHKAVLDVSEEGTEATAATTT

KFIVRSKDGPSYFTVSFNRTFLMMITNKATDGILFLGKVENPTKS

Signal peptide:
amino acids 1-20

FIGURE 137

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGC
CTGAGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAG
CCTTTATCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAG
TCTTGGTACATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAG**ATG**AAGATGC
AGAAAGGAAATGTTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCC
AATGAGACTAGCACCTCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGC
CACCAACTCTGGGTCCAGTGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCA
GCGTGACCTCCAATGGGGTCAGCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGG
ATCAGCACAGCCACCAACTCTGAGTTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAA
CTCTGAGTCCAGCACAACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAC
CCTCCAGTGGGGCCAGCACAGTCACCAACTCTGGGTCCAGTGTGACCTCCAGTGGAGCCAGC
ACTGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGA
GTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAACCTCCA
GTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGGGGCCAGCACAGCC
ACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGAGTCCAG
CACAACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGAACGACCTCCAATGGGG
CTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAAC
TCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACGAC
CTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCTAGCA
CAGCCACCAACTCTGACTCCAGCACAACCTCCAGTGGGGCCGGCACAGCCACCAACTCTGAG
TCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCTCCAG
TGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCA
CCAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGC
ACAACCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGGGGC
TAGCACAGCCACCAACTCTGACTCCAGCACAACCTCCAGTGAGGCCAGCACAGCCACCAACT
CTGAGTCTAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAACC
TCCAGTGGGGCCAACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAAC
AGCAGCTCTGACTGGAATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAGTGAGGCAA
AGCCTGGTGGGTCCCTGGTGCCGTGGGAAATCTTCCTCATCACCCTGGTCTCGGTTGTGGCG
GCCGTGGGGCTCTTTGCTGGGCTCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACAC
CTTTAACACAGCTGTCTACCACCCTCATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAG
GGAATCATGGAGCCCCCCACAGGCCCAGGTGGAGTCCTAACTGGTTCTGGAGGAGACCAGTA
TCATCGATAGCCATGGAGATGAGCGGGAGGAACAGCGGGCCC**TGA**GCAGCCCCGGAAGCAAG
TGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCACCCAAGACCTGGTTTCCTTTCATTCATC
CCAGGAGACCCCTCCCAGCTTTGTTTGAGATCCTGAAAATCTTGAAGAAGGTATTCCTCACC
TTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATATTGCTCATTTAGCTAAGAAATAA
ATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTGCCCCGGGGTGGGTATCTAG
CTCTGAGATGAACTCAGTTATAGGAGAAAACCTCCATGCTGGACTCCATCTGGCATTCAAAA
TCTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAA

FIGURE 138

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATI

SGSSVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSE

SSTPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSS

TTSSGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTT

SNGAGTATNSESSTTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSS

GASTATNSDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGA

NTATNSESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEAST

ATNSESSTVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAV

SEAKPGGSLVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGP

GPGGNHGAPHRPRWSPNWFWRRPVSSIAMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 139

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTC

CCAGCAAT**ATG**CATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCC

CTGTCTGGATGGGCGGCCAGCGATGACCCCATTGAGAAGGTCATTGAAGGGATCAACCGAGG

GCTGAGCAATGCAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGGAATCACGC

ATGCCGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGC

AAGGAGTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGAT

CAACCATGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACG

CTGCTGGACAGGCCGGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCAC

CAGGCTGGGAAGGAAGCAGAGAAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGG

AAAGGAAGTGGAGAAGCTTGGCCAAGGTGCCCACCATGCTGCTGGCCAGGCCGGGAAGGAGC

TGCAGAATGCTCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGC

AACCATCAAAGCGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGG

GGCCTCAGTCAACACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCA

TGCCC**TAA**ACTGGCATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTCACATCAGCTGAC

ATGACCTGGAGGGGTTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTTGTACTG

GGATTTGTGAATAAACTTGATACACCA

FIGURE 140

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66675

><subunit 1 of 1, 247 aa, 1 stop

><MW: 25335, pI: 7.00, NX(S/T): 0

MHLARLVGSCSLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVGKALDGINSGITHAG

REVEKVFNGLSNMGSHTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHGVNNAAG

QAGKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQN

AHNGVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWRSVANIMP

Important features of the protein:

Signal peptide:

amino acids 1-25

Homologous region to circumsporozoite (CS) repeats:

amino acids 35-225

141/270

FIGURE 141

```
CTCCGGGTCCCCAGGGGCTGCGCCGGGCCGGCCTGGCAAGGGGGACGAGTCAGTGGACACTCCAGGAAGAGCGGC
CCCGCGGGGGGGCGATGACCGTGCGCTGACCCTGACTCACTCCAGGTCCGGAGGCGGGGGCCCCCGGGGCGACTCG
GGGGCGGACCGCGGGGCGGAGCTGCCGCCCGTGAGTCCGGCCGAGCCACCTGAGCCCGAGCCGCGGGACACCGTC
GCTCCTGCTCTCCGAATGCTGCGCACCGCGATGGGCCTGAGGAGCTGGCTCGCCGCCCCATGGGGCGCGCTGCCG
CCTCGGCCACCGCTGCTGCTGCTCCTGCTGCTGCTGCTCCTGCTGCAGCCGCCGCCTCCGACCTGGGCGCTCAGC
CCCCGGATCAGCGCTGCCTCTGGGCTCTGAAGAGCGGCCATTCCTCAGATTCGAAGCTGAACACATCTCCAACTAC
ACAGCCCTTCTGCTGAGCAGGGATGGCAGGACCCTGTACGTGGGTGCTCGAGAGGCCCTCTTTGCACTCAGTAGC
AACCTCAGCTTCCTGCCAGGCGGGGAGTACCAGGAGCTGCTTTGGGGTGCAGACGCAGAGAAGAAACAGCAGTGC
AGCTTCAAGGGCAAGGACCCACAGCGCGACTGTCAAAACTACATCAAGATCCTCCTGCCGCTCAGCGGCAGTCAC
CTGTTCACCTGTGGCACAGCAGCCTTCAGCCCCATGTGTACCTACATCAACATGGAGAACTTCACCCTGGCAAGG
GACGAGAAGGGGAATGTCCTCCTGGAAGATGGCAAGGGCCGTTGTCCCTTCGACCCGAATTTCAAGTCCACTGCC
CTGGTGGTTGATGGCGAGCTCTACACTGGAACAGTCAGCAGCTTCCAAGGGAATGACCCGGCCATCTCGCGGAGC
CAAAGCCTTCGCCCCCACCAAGACCGAGAGCTCCCTCAACTGGCTGCAAGACCCAGCTTTTGTGGCCTCAGCCTAC
ATTCCTGAGAGCCTGGGCAGCTTGCAAGGCGATGATGACAAGATCTACTTTTTTCTTCAGCGAGACTGGCCAGGAA
TTTGAGTTCTTTGAGAACACCATTGTGTCCCGCATTGCCCGCATCTGCAAGGGCGATGAGGGTGGAGAGCGGGTG
CTACAGCAGCGCTGGACCTCCTTCCTCAAGGGCCAGCTGCTGTGCTCACGGCCCGACGATGGCTTCCCCTTCAAC
GTGCTGCAGGATGTCTTCACGCTGAGCCCCAGCCCCCAGGACTGGCGTGACACCCTTTTCTATGGGGTCTTCACT
TCCCAGTGGCACAGGGGAACTACAGAAGGCTCTGCCGTCTGTGTCTTCACAATGAAGGATGTGCAGAGAGTCTTC
AGCGGCCTCTACAAGGAGGTGAACCGTGAGACACAGCAGTGGTACACCGTGACCCACCCGGTGCCCACACCCCGG
CCTGGAGCGTGCATCACCAACAGTGCCCGGGAAAGGAAGATCAACTCATCCCTGCAGCTCCCAGACCGCGTGCTG
AACTTCCTCAAGGACCCACTTCCTGATGGACGGGCAGGTCCGAAGCCGCATGCTGCTGCTGCAGCCCCAGGCTCGC
TACCAGCGCGTGGCTGTACACCGCGTCCCTGGCCTGCACCACACCTACGATGTCCTCTTCCTGGGCACTGGTGAC
GGCCGGCTCCACAAGGCAGTGAGCGTGGGCCCCCGGGTGCACATCATTGAGGAGCTGCAGATCTTCTCATCGGGA
CAGCCCGTGCAGAATCTGCTCCTGGACACCCACAGGGGGCTGCTGTATGCGGCCTCACACTCGGGCGTAGTCCAG
GTGCCCATGGCCAACTGCAGCCTGTACCGGAGCTGTGGGGACTGCCTCCTCGCCCGGGACCCCTACTGTGCTTGG
AGCGGCTCCAGCTGCAAGCACGTCAGCCTCTACCAGCCTCAGCTGGCCACCAGGCCGTGGATCCAGGACATCGAG
GGAGCCAGCGCCAAGGACCTTTGCAGCGCGTCTTCGGTTGTGTCCCCGTCTTTTGTACCAACAGGGGAGAAGCCA
TGTGAGCAAGTCCAGTTCCAGCCCAACACAGTGAACACTTTGGCCTGCCCGCTCCTCTCCAACCTGGCGACCCGA
CTCTGGCTACGCAACGGGGCCCCCGTCAATGCCTCGGCCTCCTGCCACGTGCTACCCACTGGGGACCTGCTGCTG
GTGGGCACCCAACAGCTGGGGGAGTTCCAGTGCTGGTCACTAGAGGAGGGCTTCCAGCAGCTGGTAGCCAGCTAC
TGCCCAGAGGTGGTGGAGGACGGGGTGGCAGACCAAACAGATGAGGGTGGCAGTGTACCCGTCATTATCAGCACA
TCGCGTGTGAGTGCACCAGCTGGTGGCAAGGCCAGCTGGGGTGCAGACAGGTCCTACTGGAAGGAGTTCCTGGTG
ATGTGCACGCTCTTTGTGCTGGCCGTGCTGCTCCCCAGTTTTATTCTTGCTCTACCGGCACCGGAACAGCATGAAA
GTCTTCCTGAAGCAGGGGGAATGTGCCAGCGTGCACCCCAAGACCTGCCCTGTGGTGCTGCCCCCTGAGACCCGC
CCACTCAACGGCCTAGGGCCCCCTAGCACCCCGCTCGATCACCGAGGGTACCAGTCCCTGTCAGACAGCCCCCCG
GGGGCCCGAGTCTTCACTGAGTCAGAGAAGAGGCCACTCAGCATCCAAGACAGCTTCGTGGAGGTATCCCCAGTG
TGCCCCCGGCCCCGGGTCCGCCTTGGCTCGGAGATCCGTGACTCTGTGGTGTGAGAGCTGACTTCCAGAGGACGC
TGCCCTGGCTTCAGGGGCTGTGAATGCTCGGAGAGGGTCAACTGGACCTCCCCTCCGCTCTGCTCTTCGTGGAAC
ACGACCGTGGTGCCCGGCCCTTGGGAGCCTTGGAGCCAGCTGGCCTGCTGCTCTCCAGTCAAGTAGCGAAGCTCC
TACCACCCAGACACCCAAACAGCCGTGGCCCCAGAGGTCCTGGCCAAATATGGGGGCCTGCCTAGGTTGGTGGAA
CAGTGCTCCTTATGTAAACTGAGCCCTTTGTTTAAAAAAACAATTCCAAATGTGAAACTAGAATGAGAGGGAAGAG
ATAGCATGGCATGCAGCACACACGGCTGCTCCAGTTCATGGCCTCCCAGGGGTGCTGGGGATGCATCCAAAGTGG
TTGTCTGAGACAGAGTTGGAAACCCTCACCAACTGGCCTCTTCACCTTCCACATTATCCCGCTGCCACCGGCTGC
CCTGTCTCACTGCAGATTCAGGACCAGCTTGGGCTGCGTGCGTTCTGCCTTGCCAGTCAGCCGAGGATGTAGTTG
TTGCTGCCGTCGTCCCACCACCTCAGGGACCAGAGGGCTAGGTTGGCACTGCGGCCCTCACCAGGTCCTGGGCTC
GGACCCAACTCCTGGACCTTTCCAGCCTGTATCAGGCTGTGGCCACACGAGAGGACAGCGCGAGCTCAGGAGAGA
TTTCGTGACAATGTACGCCTTTCCCTCAGAATTCAGGGAAGAGACTGTCGCCTGCCTTCCTCCGTTGTTGCGTGA
GAACCCGTGTGCCCCTTCCCACCATATCCACCCTCGCTCCATCTTTGAACTCAAACACGAGGAACTAACTGCACC
CTGGTCCTCTCCCCAGTCCCCAGTTCACCCTCCATCCCTCACCTTCCTCCACTCTAAGGGATATCAACACTGCCC
AGCACAGGGGCCCTGAATTTATGTGGTTTTTATACATTTTTTAATAAGATGCACTTTATGTCATTTTTTAATAAA
GTCTGAAGAATTACTGTTTAAAAAAAAAAAA                          )
```

FIGURE 142

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA67962
><subunit 1 of 1, 837 aa, 1 stop
><MW: 92750, pI: 7.04, NX(S/T): 6
MLRTAMGLRSWLAAPWGALPPRPPLLLLLLLLLLLQPPPPTWALSPRISLPLGSEERPFLRF
EAEHISNYTALLLSRDGRTLYVGAREALFALSSNLSFLPGGEYQELLWGADAEKKQQCSFKG
KDPQRDCQNYIKILLPLSGSHLFTCGTAAFSPMCTYINMENFTLARDEKGNVLLEDGKGRCP
FDPNFKSTALVVDGELYTGTVSSFQGNDPAISRSQSLRPTKTESSLNWLQDPAFVASAYIPE
SLGSLQGDDDKIYFFFSETGQEFEFFENTIVSRIARICKGDEGGERVLQQRWTSFLKAQLLC
SRPDDGFPFNVLQDVFTLSPSPQDWRDTLFYGVFTSQWHRGTTEGSAVCVFTMKDVQRVFSG
LYKEVNRETQQWYTVTHPVPTPRPGACITNSARERKINSSLQLPDRVLNFLKDHFLMDGQVR
SRMLLLQPQARYQRVAVHRVPGLHHTYDVLFLGTGDGRLHKAVSVGPRVHIIEELQIFSSGQ
PVQNLLLDTHRGLLYAASHSGVVQVPMANCSLYRSCGDCLLARDPYCAWSGSSCKHVSLYQP
QLATRPWIQDIEGASAKDLCSASSVVSPSFVPTGEKPCEQVQFQPNTVNTLACPLLSNLATR
LWLRNGAPVNASASCHVLPTGDLLLVGTQQLGEFQCWSLEEGFQQLVASYCPEVVEDGVADQ
TDEGGSVPVIISTSRVSAPAGGKASWGADRSYWKEFLVMCTLFVLAVLLPVLFLLYRHRNSM
KVFLKQGECASVHPKTCPVVLPPETRPLNGLGPPSTPLDHRGYQSLSDSPPGARVFTESEKR
PLSIQDSFVEVSPVCPRPRVRLGSEIRDSVV
```

Transmembrane domains:
amino acids 23-46 (type II), 718-738

FIGURE 143A

```
CTAAGCCGGAGGATGTGCAGCTGCGGCGGCGGCGCCGGCTACGAAGAGGACGGGGACAGGCGCCGTGCGAACCGA
GCCCAGCCAGCCGGAGGACGCGGGCAGGGCGGGACGGGAGCCCGGACTCGTCTGCCGCCGCCGTCGTCGCCGTCG
TGCCGGCCCCGCGTCCCCGCGCGCGAGCGGGAGGAGCCGCCGCCACCTCGCGCCCGAGCCGCCGCTAGCGCGCGC
CGGGCATGGTCCCCTCTTAAAGGCGCAGGCCGCGGCGGCGGGGGCGGGTGTGCGGAACAAAGCGCCGGCGCGGGG
CCTGCGGGCGGCTCGGGGGCCGCGATGGGCGCGGCGGGCCCGCGGCGGCGGCGGCGCTGCCCGGGCCGGGCCTCG
CGGCGCTAGGGCGGGCTGGCCTCCGTGGGCGGGGGCAGCGGGCTGAGGGCGCGCGGAGCCTGCGGCGGCGGCGGC
GGCGGCGGCGGCGGCCCGGCGGGCGGAGCGGCGCGGGCATGGCCGCGCGCGGCCGGCCGGCCGCGCCTGGCTCAGCGTGC
TGCTCGGGCTCGTCCTGGGCTTCGTGCTGGCCTCGCGGCTCGTCCTGCCCCGGGCTTCCGAGCTGAAGCGAGCGG
GCCCACGGCGCCGCGCCAGCCCCGAGGGCTGCCGGTCCGGGCAGGCGGCGGCTTCCCAGGCCGGCGGGGCGCGCG
GCGATGCGCGCGGGGCGCAGCTCTGGCCGCCCGGCTCGGACCCAGATGGCGGCCCGCGCGACAGGAACTTTCTCT
TCGTGGGAGTCATGACCGCCCAGAAATACCTGCAGACTCGGGCCGTGGCCGCCTACAGAACATGGTCCAAGACAA
TTCCTGGGAAAGTTCAGTTCTTCTCAAGTGAGGGTTCTGACACATCTGTACCAATTCCAGTAGTGCCACTACGGG
GTGTGGACGACTCCTACCCGCCCCAGAAGAAGTCCTTCATGATGCTCAAGTACATGCACGACCACTACTTGGACA
AGTATGAATGGTTTATGAGAGCAGATGATGACGTGTACATCAAAGGAGACCGTCTGGAGAACTTCCTGAGGAGTT
TGAACAGCAGCGAGCCCCTCTTTCTTGGGCAGACAGGCCTGGGCACCACGGAAGAAATGGGAAAACTGGCCCTGG
AGCCTGGTGAGAACTTCTGCATGGGGGGGGCCTGGCGTGATCATGAGCCGGGAGGTGCTTCGGAGAATGGTGCCGC
ACATTGGCAAGTGTCTCCGGGAGATGTACACCACCCATGAGGACGTGGAGGTGGGAAGGTGTGTCCGGAGGTTTG
CAGGGGGTGCAGTGTGTCTGGTCTTATGAGATGCGGCAGCTTTTTTATGAGAATTACGAGCAGAACAAAAAGGGGT
ACATTAGAGATCTCCATAACAGTAAAATTCACCAAGCTATCACATTACACCCCAACAAAAACCCACCCTACCAGT
ACAGGCTCCACAGCTACATGCTGAGCCGCAAGATATCCGAGCTCCGCCATCGCACAATACAGCTGCACCGCGAAA
TTGTCCTGATGAGCAAATACAGCAACACAGAAATTCATAAAGAGGACCTCCAGCTGGGAATCCCTCCCTCCTTCA
TGAGGTTTCAGCCCCGCCAGCGAGAGGAGATTCTGGAATGGGAGTTTCTGACTGGAAAATACTTGTATTCGGCAG
TTGACGGCCAGCCCCCTCGAAGAGGAATGGACTCCGCCCAGAGGGAAGCCTTGGACGACATTGTCATGCAGGTCA
TGGAGATGATCAATGCCAACGCCAAGACCAGAGGGCGCATCATTGACTTCAAAGAGATCCAGTACGGCTACCGCC
GGGTGAACCCCATGTATGGGGCTGAGTACATCCTGGACCTGCTGCTTCTGTACAAAAAGCACAAAGGGAAGAAAA
TGACGGTCCCTGTGAGGAGGCACGCGTATTTACAGCAGACTTTCAGCAAAATCCAGTTTGTGGAGCATGAGGAGC
TGGATGCACAAGAGTTGGCCAAGAGAATCAATCAGGAATCTGGATCCTTGTCCTTTCTCTCAAACTCCCTGAAGA
AGCTCGTCCCCTTTCAGCTCCCTGGGTCGAAGAGTGAGCACAAAGAACCCAAAGATAAAAAGATAAACATACTGA
TTCCTTTGTCTGGGCGTTTCGACATGTTTGTGAGATTTATGGGAAACTTTGAGAAGACGTGTCTTATCCCCAATC
AGAACGTCAAGCTCGTGGTTCTGCTTTTCAATTCTGACTCCAACCCTGACAAGGCCAAACAAGTTGAACTGATGA
GAGATTACCGCATTAAGTACCCTAAAGCCGACATGCAGATTTTGCCTGTGTCTGGAGAGTTTTCAAGAGCCCTGG
CCCTGGAAGTAGGATCCTCCCAGTTTAACAATGAATCTTTGCTCTTCTTCTGCGACGTCGACCCTCGTGTTTACTA
CAGAATTCCTTCAGCGATGTCGAGCAAATACAGTTCTGGGCCAACAAATATATTTTCCAATCATCTTCAGCCAGT
ATGACCCAAAGATTGTTTATAGTGGGAAAGTTCCCAGTGACAACCATTTTGCCTTTACTCAGAAAACTGGCTTCT
GGAGAAACTATGGGTTTGGCATCACGTGTATTTATAAGGGAGATCTTGTCCGAGTGGGTGGCTTTGATGTTTCCA
TCCAAGGCTGGGGGGCTGGAGGATGTGGACCTTTTCAACAAGGTTGTCCAGGCAGGTTTGAAGACGTTTAGGAGCC
AGGAAGTAGGAGTAGTCCACGTCCACCATCCTGTCTTTTGTGATCCCAATCTTGACCCCAAACAGTACAAAATGT
GCTTGGGGTCCAAAGCATCGACCTATGGGTCCACCCAGCAGCTGGCTGAGATGTGGCTGGAAAAAAAATGATCCAA
GTTACAGTAAAAGCAGCAATAATAATGGCTCAGTGAGGACAGCCTAATGTCCAGCTTTGCTGGAAAAGACGTTTT
TAATTATCTAATTTATTTTTCAAAAATTTTTTGTATGATCAGTTTTTGAAGTCCGTATACAAGGATATATTTTAC
AAGTGGTTTTCTTACATAGGACTCCTTTAAGATTGAGCTTTCTGAACAAGAAGGTGATCAGTGTTTGCCTTTGAA
CACATCTTCTTGCTGAACATTATGTAGCAGACCTGCTTAACTTTGACTTGAAATGTACCTGATGAACAAAACTTT
TTTAAAAAAATGTTTTCTTTTGAGACCCTTTGCTCCAGTCCTATGGCAGAAAACGTGAACATTCCTGCAAAGTAT
TATTGTAACAAAACACTGTAACTCTGGTAAATGTTCTGTTGTGATTGTTAACATTCCACAGATTCTACCTTTTGT
GTTTTGTTTTTTTTTTTACAATTGTTTTAAAGCCATTTCATGTTCCAGTTGTAAGATAAGGAAATGTGATAATA
GCTGTTTCATCATTGTCTTCAGGAGAGCTTTCCAGAGTTGATCATTTCCTCTCATGGTACTCTGCTCAGCATGGC
CACGTAGGTTTTTTGTTTGTTTTGTTTTGTTCTTTTTTTGAGACGGAGTCTCACTCTGTTACCCAGGCTGGAATG
CAGTGGCGCAATCTTGGCTCACTTTAACCTCCACTTCCCTGGTTCAAGCAATTCCCCTGCCTTTCCTCCCGAGT
AGCTGGGATTACAGGCACACACCACCACGCCCAGNTAGTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCAT
GCAAGCCCAGCTGGCCACGTAGGTTTTAAAGCAAGGGGCGTGAAGAAGGCACAGTGAGGTATGTGGCTGTTCTCG
TGGTAGTTCATTCGGCCTAAATAGACCTGGCATTAAATTTCAAGAAGGATTTGGCATTTTCTCTTCTTGACCCTT
CTCTTTAAAGGGTAAAATATTAATGTTTAGAATGACAAAGATGAATTATTACAATAAATCTGATGTACACAGACT
GAAACATACACACATACACCCTAATCAAAACGTTGGGGAAAAATGTATTTGGTTTTGTTCCTTTCATCCTGTCTG
TGTTATGTGGGTGGAGATGGTTTTCATTCTTTCATTACTGTTTTGTTTTATCCTTTGTATCTGAAATACCTTTAA
TTTATTTAATATCTGTTGTTCAGAGCTCTGCCATTTCTTGAGTACCTGTTAGTTAGTATTATTTATGTGTATCGG
GAGTGTGTTTAGTCTGTTTTATTTGCAGTAAACCGATCTCCAAAGATTTCCTTTTGGAAACGCTTTTTCCCCTCC
```

144/270

FIGURE 143B

TTAATTTTTATATTCCTTACTGTTTTACTAAATATTAAGTGTTCTTTGACAATTTTGGTGCTCATGTGTTTTGGG
GACAAAAGTGAAATGAATCTGTCATTATACCAGAAAGTTAAATTCTCAGATCAAATGTGCCTTAATAAATTTGTT
TTCATTTAGATTTCAAACAGTGATAGACTTGCCATTTTAATACACGTCATTGGAGGGCTGCGTATTTGTAAATAG
CCTGATGCTCATTTGGAAAAATAAACCAGTGAACAATATTTTTCTATTGTACTTTTCGAACCATTTTGTCTCATT
ATTCCTGTTTTAGCTGAAGAATTGTATTACATTTGGAGAGTAAAAAACTTAAACACGAAAAAA

FIGURE 144

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68836
><subunit 1 of 1, 802 aa, 1 stop
><MW: 91812, pI: 9.52, NX(S/T): 3
MAARGRRAWLSVLLGLVLGFVLASRLVLPRASELKRAGPRRRASPEGCRSGQAAASQAGGAR
GDARGAQLWPPGSDPDGGPRDRNFLFVGVMTAQKYLQTRAVAAYRTWSKTIPGKVQFFSSEG
SDTSVPIPVVPLRGVDDSYPPQKKSFMMLKYMHDHYLDKYEWFMRADDDVYIKGDRLENFLR
SLNSSEPLFLGQTGLGTTEEMGKLALEPGENFCMGGPGVIMSREVLRRMVPHIGKCLREMYT
THEDVEVGRCVRRFAGVQCVWSYEMRQLFYENYEQNKKGYIRDLHNSKIHQAITLHPNKNPP
YQYRLHSYMLSRKISELRHRTIQLHREIVLMSKYSNTEIHKEDLQLGIPPSFMRFQPRQREE
ILEWEFLTGKYLYSAVDGQPPRRGMDSAQREALDDIVMQVMEMINANAKTRGRIIDFKEIQY
GYRRVNPMYGAEYILDLLLLYKKHKGKKMTVPVRRHAYLQQTFSKIQFVEHEELDAQELAKR
INQESGSLSFLSNSLKKLVPFQLPGSKSEHKEPKDKKINILIPLSGRFDMFVRFMGNFEKTC
LIPNQNVKLVVLLFNSDSNPDKAKQVELMRDYRIKYPKADMQILPVSGEFSRALALEVGSSQ
FNNESLLFFCDVDLVFTTEFLQRCRANTVLGQQIYFPIIFSQYDPKIVYSGKVPSDNHFAFT
QKTGFWRNYGFGITCIYKGDLVRVGGFDVSIQGWGLEDVDLFNKVVQAGLKTFRSQEVGVVH
VHHPVFCDPNLDPKQYKMCLGSKASTYGSTQQLAEMWLEKNDPSYSKSSNNNGSVRTA
```

Signal peptide:
amino acids 1-23

146/276

FIGURE 145

GGACAACCGTTGCTGGGTGTCCCAGGGCCTGAGGCAGGACGGTACTCCGCTGACACCTTCCC

TTTCGGCCTTGAGGTTCCCAGCCTGGTGGCCCCAGGACGTTCCGGTCGCATGGCAGAGTGCT

ACGGACGACGCCT**ATG**AAGCCCTTAGTCCTTCTAGTTGCGCTTTTGCTATGGCCTTCGTCTG

TGCCGGCTTATCCGAGCATAACTGTGACACCTGATGAAGAGCAAAACTTGAATCATTATATA

CAAGTTTTAGAGAACCTAGTACGAAGTGTTCCCTCTGGGGAGCCAGGTCGTGAGAAAAAATC

TAACTCTCCAAAACATGTTTATTCTATAGCATCAAAGGGATCAAAATTTAAGGAGCTAGTTA

CACATGGAGACGCTTCAACTGAGAATGATGTTTTAACCAATCCTATCAGTGAAGAAACTACA

ACTTTCCCTACAGGAGGCTTCACACCGGAAATAGGAAAGAAAAAACACACGGAAAGTACCCC

ATTCTGGTCGATCAAACCAAACAATGTTTCCATTGTTTTGCATGCAGAGGAACCTTATATTG

AAAATGAAGAGCCAGAGCCAGAGCCGGAGCCAGCTGCAAAACAAACTGAGGCACCAAGAATG

TTGCCAGTTGTTACTGAATCATCTACAAGTCCATATGTTACCTCATACAAGTCACCTGTCAC

CACTTTAGATAAGAGCACTGGCATTGAGATCTCTACAGAATCAGAAGATGTTCCTCAGCTCT

CAGGTGAAACTGCGATAGAAAAACCCGAAGAGTTTGGAAAGCACCCAGAGAGTTGGAATAAT

GATGACATTTTGAAAAAAATTTTAGATATTAATTCACAAGTGCAACAGGCACTTCTTAGTGA

CACCAGCAACCCAGCATATAGAGAAGATATTGAAGCCTCTAAAGATCACCTAAAACGAAGCC

TTGCTCTAGCAGCAGCAGCAGAACATAAATTAAAAACAATGTATAAGTCCCAGTTATTGCCA

GTAGGACGAACAAGTAATAAAATTGATGACATCGAAACTGTTATTAACATGCTGTGTAATTC

TAGATCTAAACTCTATGAATATTTAGATATTAAATGTGTTCCACCAGAGATGAGAGAAAAAG

CTGCTACAGTATTCAATACATTAAAAAAATATGTGTAGATCAAGGAGAGTCACAGCCTTATTA

AAAGTTTAT**TAA**CAATAATATAAAAATTTTAAACCTACTTGATATTCCATAACAAAGCTGA

TTTAAGCAAACTGCATTTTTTCACAGGAGAAATAATCATATTCGTAATTTCAAAAGTTGTAT

AAAAATATTTTCTATTGTAGTTCAAATGTGCCAACATCTTTATGTGTCATGTGTTATGAACA

ATTTTCATATGCACTAAAAACCTAATTTAAAATAAAATTTTGGTTCAGGAAAAAA

147/270

FIGURE 146

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68864

><subunit 1 of 1, 350 aa, 1 stop

><MW: 39003, pI: 5.59, NX(S/T): 1

MKPLVLLVALLLWPSSVPAYPSITVTPDEEQNLNHYIQVLENLVRSVPSGEPGREKKSNSPK

HVYSIASKGSKFKELVTHGDASTENDVLTNPISEETTTFPTGGFTPEIGKKKHTESTPFWSI

KPNNVSIVLHAEEPYIENEEPEPEPEPAAKQTEAPRMLPVVTESSTSPYVTSYKSPVTTLDK

STGIEISTESEDVPQLSGETAIEKPEEFGKHPESWNNDDILKKILDINSQVQQALLSDTSNP

AYREDIEASKDHLKRSLALAAAAEHKLKTMYKSQLLPVGRTSNKIDDIETVINMLCNSRSKL

YEYLDIKCVPPEMREKAATVFNTLKNMCRSRRVTALLKVY

Signal peptide:

amino acids 1-19

FIGURE 147

```
CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTG
ACCTCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTCATAGTGTGAGATCAACCCACA
GGAATATCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTCAGGACA
GTGGCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCT
CCTGCTCCCTCTTTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAG
TTCCATGCTGTGGTCCACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACA
GTATCGAGGGAGAACTGAGTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGC
TAAAAAACATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTCAGTTCCCAGATTTAC
GATGAGGAGGCCACCTGGGAGCTGCGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCAT
CGTGGGATATGTTGACGGAGGTATCCAGTTACTCTGCCTGTCCTCAGGCTGGTTCCCCCAGC
CCACAGCCAAGTGGAAAGGTCCACAAGGACAGGATTTGTCTTCAGACTCCAGAGCAAATGCA
GATGGGTACAGCCTGTATGATGTGGAGATCTCCATTATAGTCCAGGAAATGCTGGGAGCAT
ATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAG
AGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGT
GCCCTGTGTGGTGTTGTCATGGGGATGATAATTGTTTTCTTCAAATCCAAAGGGAAAATCCA
GGCGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACG
CAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAA
ACTGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAA
GAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGGAC
AAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACAATGTG
ACTTTGTCTCCCAACAATGGGTATTGGGTCCTCAGACTGACAACAGAACATTTGTATTTCAC
ATTCAATCCCCATTTTATCAGCCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCC
TGGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACC
CTGCTGACATGTCAGTTTGAAGGCTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGA
GGAAAAGGGGACTCCCATATTCATATGTCCAGTGTCCTGGGGATGAGACAGAGAAGACCCTG
CTTAAAGGGCCCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGC
CCCAGCTTCCTCTCCGGAGCCTGCGCACAGAGAGTCACGCCCCCCACTCTCCTTTAGGGAGC
TGAGGTTCTTCTGCCCTGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGAT
TGGCCTGACCCTGTGGGAGTCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCA
CATTAGGTTTAGTTTGTGAAAACTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCC
CAGGCTCCTCATTTGCTAGTCACGGACAGTGATTCCTGCCTCACAGGTGAAGATTAAAGAGA
CAACGAATGTGAATCATGCTTGCAGGTTTGAGGGCACAGTGTTTGCTAATGATGTGTTTTTA
TATTATACATTTTCCCACCATAAACTCTGTTTGCTTATTCCACATTAATTTACTTTTCTCTA
TACCAAATCACCCATGGAATAGTTATTGAACACCTGCTTTGTGAGGCTCAAAGAATAAAGAG
GAGGTAGGATTTTTCACTGATTCTATAAGCCCAGCATTACCTGATACCAAAACCAGGCAAAG
AAAACAGAAGAAGAGGAAGGAAAACTACAGGTCCATATCCCTCATTAACACAGACACAAAAA
TTCTAAATAAAATTTTAACAAATTAAACTAAACAATATATTTAAAGATGATATATAACTACT
CAGTGTGGTTTGTCCCACAAATGCAGAGTTGGTTTAATATTTAAATATCAACCAGTGTAATT
CAGCACATTAATAAAGTAAAAAAGAAACCATAAAAAAAAAAAAAAAA
```

FIGURE 148

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68866
><subunit 1 of 1, 466 aa, 1 stop
><MW: 52279, pI: 6.16, NX(S/T): 2
MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHA
VVHLYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEE
ATWELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGY
SLYDVEISIIVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALC
GVVMGMIIVFFKSKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVT
HRKAPQEVPHSEKRFTRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLS
PNNGYWVLRLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLT
CQFEGLLRPYIQHAMYDEEKGTPIFICPVSWG
```

Signal peptide:
amino acids 1-17

Transmembrane domains:
amino acids 131-150, 235-259

FIGURE 149

```
CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTA
GGAAAAGAGTTTGTTGGGAACCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTC
CTGGCAGTGTGCATTGGACTCACTGTTCATTATGTGAGATATAATCAAAAGAAGACCTACAA
TTACTATAGCACATTGTCATTTACAACTGACAAACTATATGCTGAGTTTGGCAGAGAGGCTT
CTAACAATTTTACAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAA
TCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGG
AGTGTTGGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAAACTGTAG
ATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCTAAAGTA
GATCCTCACTCAGTTAAAATTAAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCA
TTGCTGCGGAACACGAAGAAGTAAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAG
AAGTAGAAGAGGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGT
GGAGCAACCTTAATTAATGCCACATGGCTTGTGAGTGCTGCTCACTGTTTTACAACATATAA
GAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACAATAAAACCTTCGAAAATGAAACGGG
GTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACATGACTATGATATTTCT
CTTGCAGAGCTTTCTAGCCCTGTTCCCTACACACAAATGCAGTACATAGAGTTTGTCTCCCTGA
TGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAA
ATGATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACT
TGCAATGAACCTCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTT
AGAAGGAAAAACAGATGCATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTA
GAGATATCTGGTACCTTGCTGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCCAACAAG
CCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTGGATTACTTCAAAAACTGGTATCTA
AGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTTGGGTGTGGAGGCCATT
TTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAAC
TGTTTGCTTGATGCATGTATTTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTG
CCAGATCAACTCTGTCATCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATA
ATACAATATTACATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTTGTCAGAATTTTGAC
TTGTTGACATAAATTTGTAATGCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTC
AGCTCCTCTCATTTCAGCAAATATCCATTTTCAAGGTGCAGAACAAGGAGTGAAAGAAAATA
TAAGAAGAAAAAAATCCCCTACATTTTATTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGT
GGAATATTAGAAATGATCATATTCATTATGAAAGGTCAAGCAAAGACAGCAGAATACCAATC
ACTTCATCATTTAGGAAGTATGGGAACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAGATATA
TCCTTATTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTG
ACCTATAATAATTATACAAACTTCATGCAATGTACTTGTTCTAAGCAAATTAAAGCAAATAT
TTATTTAACATTGTTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA
```

FIGURE 150

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68871

><subunit 1 of 1, 423 aa, 1 stop

><MW: 47696, pI: 8.96, NX(S/T): 3

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTD
KLYAEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICR
FHSTEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTL
GQSLRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFG
VTIKPSKMKRGLRRIIVHEKYKHPSHDYDISLAELSSPVPYTNAVHRVCLPDASYEFQPGDV
MFVTGFGALKNDGYSQNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGD
SGGPLVSSDARDIWYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 151

GTCGAAGGTTATAAAAGCTTCCAGCCAAACGGCATTGAAGTTGAAGATACAACCTGACAGCA

CAGCCTGAGATCTTGGGGATCCCTCAGCCTAACACCCACAGACGTCAGCTGGTGGATTCCCG

CTGCATCAAGGCCTACCCACTGTCTCC**ATG**CTGGGCTCTCCCTGCCTTCTGTGGCTCCTGGC

CGTGACCTTCTTGGTTCCCAGAGCTCAGCCCTTGGCCCCTCAAGACTTTGAAGAAGAGGAGG

CAGATGAGACTGAGACGGCGTGGCCGCCTTTGCCGGCTGTCCCCTGCGACTACGACCACTGC

CGACACCTGCAGGTGCCCTGCAAGGAGCTACAGAGGGTCGGGCCGGCGGCCTGCCTGTGCCC

AGGACTCTCCAGCCCCGCCCAGCCGCCCGACCCGCCGCGCATGGGAGAAGTGCGCATTGCGG

CCGAAGAGGGCCGCGCAGTGGTCCACTGGTGTGCCCCCTTCTCCCCGGTCCTCCACTACTGG

CTGCTGCTTTGGGACGGCAGCGAGGCTGCGCAGAAGGGGCCCCCGCTGAACGCTACGGTCCG

CAGAGCCGAACTGAAGGGGCTGAAGCCAGGGGGCATTTATGTCGTTTGCGTAGTGGCCGCTA

ACGAGGCCGGGGCAAGCCGCGTGCCCCAGGCTGGAGGAGAGGGCCTCGAGGGGGCCGACATC

CCTGCCTTCGGGCCTTGCAGCCGCCTTGCGGTGCCGCCCAACCCCCGCACTCTGGTCCACGC

GGCCGTCGGGGTGGGCACGGCCCTGGCCCTGCTAAGCTGTGCCGCCCTGGTGTGGCACTTCT

GCCTGCGCGATCGCTGGGGCTGCCCGCGCCGAGCCGCCGCCCGAGCCGCAGGGGCGCTC**TGA**

AAGGGGCCTGGGGGCATCTCGGGCACAGACAGCCCCACCTGGGGCGCTCAGCCTGGCCCCCG

GGAAAGAGGAAAACCCGCTGCCTCCAGGGAGGGCTGGACGGCGAGCTGGGAGCCAGCCCCAG

GCTCCAGGGCCACGGCGGAGTCATGGTTCTCAGGACTGAGCGCTTGTTTAGGTCCGGTACTT

GGCGCTTTGTTTCCTGGCTGAGGTCTGGGAAGGAATAGAAAGGGGCCCCCAATTTTTTTTTA

AGCGGCCAGATAATAAATAATGTAACCTTTGCGGTTAAAAAAAAAAAAAAAAAAAA

FIGURE 152

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68874

><subunit 1 of 1, 238 aa, 1 stop

><MW: 25262, pI: 6.44, NX(S/T): 1

MLGSPCLLWLLAVTFLVPRAQPLAPQDFEEEEADETETAWPPLPAVPCDYDHCRHLQVPCKE

LQRVGPAACLCPGLSSPAQPPDPPRMGEVRIAAEEGRAVVHWCAPFSPVLHYWLLLWDGSEA

AQKGPPLNATVRRAELKGLKPGGIYVVCVVAANEAGASRVPQAGGEGLEGADIPAFGPCSRL

AVPPNPRTLVHAAVGVGTALALLSCAALVWHFCLRDRWGCPRRAAARAAGAL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 194-220

N-glycosylation site.

amino acids 132-135

154/270

FIGURE 153

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCC
CTGCCCCGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGG
CCCAGCGCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGC
TTCTCCCTTACGGGGCTCACA<u>ATG</u>GCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCT
CTACGCCCTCAATCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGA
TGAGGGACTACCTAAATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTC
ATTTTGACTTACTTTCCTGTGGTTCATCCGGTCATGATTGCTGTTTGCTGTTTCCTTATCAT
TGTGGGGATGTTAGGATATTGTGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACT
TTGGAAGTTTGCTTGTCATTTTCTGTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAG
GAACTTATGGTTCCAGTACAATGGTCAGATATGGTCACTTTGAAAGCCAGGATGACAAATTA
TGGATTACCTAGATATCGGTGGCTTACTCATGCTTGGAATTTTTTTCAGAGAGAGTTTAAGT
GCTGTGGAGTAGTATATTTCACTGACTGGTTGGAAATGACAGAGATGGACTGGCCCCCAGAT
TCCTGCTGTGTTAGAGAATTCCCAGGATGTTCCAAACAGGCCCACCAGGAAGATCTCAGTGA
CCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTTTTTGAGAGGAACCAAACAACTGC
AGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAATCCTGGCCATGA<u>t</u>TCTCACC
ATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGGACAGACCAAATGATGTC
CTTGAAGAATGACAACTCTCAGCACCTGTCATGTCCCTCAGTAGAACTGTTGAAACCAAGCC
TGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACTTTGAGATGGAG
GAGTTA<u>TAA</u>AAAGAAATGTCACAGAAGAAAACCACAAACTTGTTTTATTGGACTTGTGAATT
TTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAATAACA
CCTAAGCATATACTATTCTATGCTTTAAAATGAGGATGGAAAAGTTTCATGTCATAAGTCAC
CACCTGGACAATAATTGATGCCCTTAAAATGCTGAAGACAGATGTCATACCCACTGTGTAGC
CTGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTC
CGCATCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTT
CTACCAACTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATA
ACTTTTATTACTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACTTTC
AATATTGGTGACTACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCACTTAAAA
GAGCAAGCTAACACATTGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTTAAA
TCTGTATAATTCAGTCGATTTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAA
ATTTGTCCTGTATAGCATCATTATTTTTAGCCTTTCCTGTTAATAAAGCTTTACTATTCTGT
CCTGGGCTTATATTACACATATAACTGTTATTTAAATACTTAACCACTAATTTTGAAAATTA
CCAGTGTGATACATAGGAATCATTATTCAGAATGTAGTCTGGTCTTTAGGAAGTATTAATAA
GAAAATTTGCACATAACTTAGTTGATTCAGAAAGGACTTGTATGCTGTTTTTCTCCCAAATG
AAGACTCTTTTTGACACTAAACACTTTTTAAAAAGCTTATCTTTGCCTTCTCCAAACAAGAA
GCAATAGTCTCCAAGTCAATATAAATTCTACAGAAAATAGTGTTCTTTTTCTCCAGAAAAAT
GCTTGTGAGAATCATTAAAAACATGTGACAATTTAGAGATTCTTTGTTTTATTTCACTGATTA
ATATACTGTGGCAAATTACACAGATTATTAAATTTTTTTACAAGAGTATAGTATATTTATTT
GAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTTTATTTCTCAGAATATGGAA
AGAAAATTAAAATGTGTCAATAAATATTTTCTAGAGAGTAA

FIGURE 154

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68880

><subunit 1 of 1, 305 aa, 1 stop

><MW: 35383, pI: 5.99, NX(S/T): 0

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTLTAETRVEEAVILTYFPV

VHPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQ

WSDMVTLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREF

PGCSKQAHQEDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWAL

YYDRREPGTDQMMSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

156/270

FIGURE 155

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCC
TGCACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGA
CCTGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAG
ACCGGGAGGATCACAGAGCCAGC**ATG**TTACAGGATCCTGACAGTGATCAACCTCTGAACAGC
CTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGAT
CCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCCTCATCAAGG
TGATTCTGGATAAATACTACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAG
CTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTT
CCCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGG
ACTCGGCCACAGGGAACTGGTTCTCTGCCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAG
ACAGCCTGTAGGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGA
TGTTGTTGAAATCACAGAAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTC
TCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCTGTGGGAAGAGCCTGAAGACCCCCCGT
GTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGGTCAGCATCCAGTACGA
CAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCCTCACGGCAGCCCACT
GCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGC
AGCTTCCCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAA
AGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGC
CCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGA
TGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCA
GGTCATTGACAGCACACGGTGCAATGCAGACGATGCGTACCAGGGGGAAGTCACCGAGAAGA
TGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCC
CTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGG
GGGCCCGAGCACCCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATG
TCTGGAAGGCTGAGCTG**TAA**TGCTGCTGCCCCTTTGCAGTGCTGGGAGCCGCTTCCTTCCTG
CCCTGCCCACCTGGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCC
CTCTGCCCACAGCCTCAGCATTTCTTGGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCC
TCGCAGCCCAGAGGCGCCCAGAGGAAGTCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCC
AGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAA
GGAACTTTCCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAAGCCCAGATCACTGTGGG
CTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTCACCCATCCCCAAGCCTA
CTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTT
ACCTACTGTTGTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCT
CTGGCAAAAAAAAAAAAA

FIGURE 156

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68885

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47644, pI: 5.18, NX(S/T): 2

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLIKVILDKYYF

LCGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWF

SACFDNFTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELRMRNSSGPCLSGSLVSL

HCLACGKSLKTPRVVGGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDV

FNWKVRAGSDKLGSFPSLAVAKIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFD

EELTPATPLWIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPE

GGVDTCQGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 157

```
GGGCTGAGGCACTGAGAGACCGGAAAGCCTGGCATTCCAGAGGGAGGGAAACGCAGCGGCATCCCCAGGCTCCAG
AGCTCCCTGGTGACAGTCTGTGGCTGAGCATGCCCTCCCAGCCCTGGGCCTGGACCCCTGGAGCCTCCTGGGCC
TTTTCCTCTTCCAACTGCTTCAGCTGCTGCTGCCGACGACGACCGCGGGGGGAGGCGGGCAGGGGCCCATGCCCA
GGGTCAGATACTATGCAGGGGATGAACGTAGGGCACTTAGCTTCTTCCACCAGAAGGGCCTCCAGGATTTTGACA
CTCTGCTCCTGAGTGGTGATGGAAATACTCTCTACGTGGGGGCTCGAGAAGCCATTCTGGCCTTGGATATCCAGG
ATCCAGGGGTCCCCAGGCTAAAGAACATGATACCGTGGCCAGCCAGTGACAGAAAAAAGAGTGAATGTGCCTTTA
AGAAGAAGAGCAATGAGACACAGTGTTTCAACTTCATCCGTGTCCTGGTTTCTTACAATGTCACCCATCTCTACA
CCTGCGGCACCTTCGCCTTCAGCCCTGCTTGTACCTTCATTGAACTTCAAGATTCCTACCTGTTGCCCATCTCGG
AGGACAAGGTCATGGAGGGAAAAGGCCAAAGCCCCTTTGACCCCGCTCACAAGCATACGGCTGTCTTGGTGGATG
GGATGCTCTATTCTGGTACTATGAACAACTTCCTGGGCAGTGAGCCCATCCTGATGCGCACACTGGGATCCCAGC
CTGTCCTCAAGACCGACAACTTCCTCCGCTGGCTGCATCATGACGCCTCCTTTGTGGCAGCCATCCCTTCGACCC
AGGTCGTCTACTTCTTCTTCGAGGAGACAGCCAGCGAGTTTGACTTCTTTGAGAGGCTCCACACATCGCGGGTGG
CTAGAGTCTGCAAGAATGACGTGGGCGGCGAAAAGCTGCTGCAGAAGAAGTGGACCACCTTCCTGAAGGCCCAGC
TGCTCTGCACCCAGCCGGGGCAGCTGCCCCTTCAACGTCATCCGCCACGCGGTCCTGCTCCCCGCCGATTCTCCCA
CAGCTCCCCACATCTACGCAGTCTTCACCTCCCAGTGGCAGGTTGGCGGGACCAGGAGCTCTGCGGTTTGTGCCT
TCTCTCTCTTGGACATTGAACGTGTCTTTAAGGGGAAATACAAAGAGTTGAACAAAGAAACTTCACGCTGGACTA
CTTATAGGGGCCCTGAGACCAACCCCCGGCCAGGCAGTTGCTCAGTGGGCCCCTCCTCTGATAAGGCCCTGACCT
TCATGAAGGACCATTTCCTGATGGATGAGCAAGTGGTGGGGACGCCCCTGCTGGTGAAATCTGGCGTGGAGTATA
CACGGCTTGCAGTGGAGACAGCCCAGGGCCTTGATGGGCACAGCCATCTTGTCATGTACCTGGGAACCACCACAG
GGTCGCTCCACAAGGCTGTGGTAAGTGGGGACAGCAGTGCTCATCTGGTGGAAGAGATTCAGCTGTTCCCTGACC
CTGAACCTGTTCGCAACCTGCAGCTGGCCCCCACCCAGGGTGCAGTGTTTGTAGGCTTCTCAGGAGGTGTCTGGA
GGGTGCCCCGAGCCAACTGTAGTGTCTATGAGAGCTGTGTGGACTGTGTCCTTGCCCGGGACCCCCACTGTGCCT
GGGACCCTGAGTCCCGAACCTGTTGCCTCCTGTCTGCCCCCAACCTGAACTCCTGGAAGCAGGACATGGAGCGGG
GGAACCCAGAGTGGGCATGTGCCAGTGGCCCCATGAGCAGGAGCCTTCGGCCTCAGAGCCGCCCGCAAATCATTA
AAGAAGTCCTGGCTGTCCCCAACTCCATCCTGGAGCTCCCCTGCCCCCACCTGTCAGCCTTGGCCTCTTATTATT
GGGAGTCATGGCCCAGCAGCAGTCCCAGAAGCCTCTTCCACTGTCTACAATGGCTCCCTCTTGCTGATAGTGCAGG
ATGGAGTTGGGGGTCTCTACCAGTGCTGGGCAACTGAGAATGGCTTTTCATACCCTGTGATCTCCTACTGGGTGG
ACAGCCAGGACCAGACCCTGGCCCTGGATCCTGAACTGGCAGGCATCCCCCGGGAGCATGTGAAGGTCCCGTTGA
CCAGGGTCAGTGGTGGGGCCGCCCTGGCTGCCCAGCAGTCCTACTGGCCCCACTTTGTCACTGTCACTGTCCTCT
TTGCCTTAGTGCTTTCAGGAGCCCTCATCATCCTCGTGGCCTCCCCATTGAGAGCACTCCGGGCTCGGGGCCAAGG
TTCAGGGCTGTGAGACCCTGCGCCCTGGGGAGAAGGCCCCGTTAAGCAGAGAGCAACACCTCCAGTCTCCCAAGG
AATGCAGGACCTCTGCCAGTGATGTGGACGCTGACAACAACTGCCTAGGCACTGAGGTAGCTTAAACTCTAGGCA
CAGGCCGGGGCTGCGGTGCAGGCACCTGGCCATGCTGGCTGGGCGGCCCAAGCACAGCCCTGACTAGGATGACAG
CAGCACAAAAGACCACCTTTCTCCCCTGAGAGGAGCTTCTGCTACTCTGCATCACTGATGACACTCAGCAGGGTG
ATGCACAGCAGTCTGCCTCCCCTATGGGACTCCCTTCTACCAAGCACATGAGCTCTCTAACAGGGTGGGGGCTAC
CCCCAGACCTGCTCCTACACTGATATTGAAGAACCTGGAGAGGATCCTTCAGTTCTGGCCATTCCAGGGACCCTC
CAGAAACACAGTGTTTCAAGAGACCCTAAAAAAACCTGCCTGTCCCAGGACCCTATGGTAATGAACACCAAACATC
TAAACAATCATATGCTAACATGCCACTCCTGGAAACTCCACTCTGAAGCTGCCGCTTTGGACACCAACACTCCCT
TCTCCCAGGGTCATGCAGGGATCTGCTCCCTCCTGCTTCCCTTACCAGTCGTGCACCGCTGACTCCCAGGAAGTC
TTTCCTGAAGTCTGACCACCTTTCTTCTTGCTTCAGTTGGGGCAGACTCTGATCCCTTCTGCCCTGGCAGAATGG
CAGGGGTAATCTGAGCCTTCTTCACTCCTTTACCCTAGCTGACCCCTTCACCTCTCCCCCTCCCTTTTCCTTTGT
TTTGGGATTCAGAAAACTGCTTGTCAGAGACTGTTTATTTTTTATTAAAAAATATAAGGCTTAAAAAAA
```

FIGURE 158

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71166

><subunit 1 of 1, 761 aa, 1 stop

><MW: 83574, pI: 6.78, NX(S/T): 4

MALPALGLDPWSLLGLFLFQLLQLLLPTTTAGGGGQGPMPRVRYYAGDERRALSFFHQKGLQ

DFDTLLLSGDGNTLYVGAREAILALDIQDPGVPRLKNMIPWPASDRKKSECAFKKKSNETQC

FNFIRVLVSYNVTHLYTCGTFAFSPACTFIELQDSYLLPISEDKVMEGKGQSPFDPAHKHTA

VLVDGMLYSGTMNNFLGSEPILMRTLGSQPVLKTDNFLRWLHHDASFVAAIPSTQVVYFFFE

ETASEFDFFERLHTSRVARVCKNDVGGEKLLQKKWTTFLKAQLLCTQPGQLPFNVIRHAVLL

PADSPTAPHIYAVFTSQWQVGGTRSSAVCAFSLLDIERVFKGKYKELNKETSRWTTYRGPET

NPRPGSCSVGPSSDKALTFMKDHFLMDEQVVGTPLLVKSGVEYTRLAVETAQGLDGHSHLVM

YLGTTTGSLHKAVVSGDSSAHLVEEIQLFPDPEPVRNLQLAPTQGAVFVGFSGGVWRVPRAN

CSVYESCVDCVLARDPHCAWDPESRTCCLLSAPNLNSWKQDMERGNPEWACASGPMSRSLRP

QSRPQIIKEVLAVPNSILELPCPHLSALASYYWSHGPAAVPEASSTVYNGSLLLIVQDGVGG

LYQCWATENGFSYPVISYWVDSQDQTLALDPELAGIPREHVKVPLTRVSGGAALAAQQSYWP

HFVTVTVLFALVLSGALIILVASPLRALRARGKVQGCETLRPGEKAPLSREQHLQSPKECRT

SASDVDADNNCLGTEVA

Signal peptide:

amino acids 1-30

Transmembrane domains:

amino acids 136-156, 222-247, 474-490, 685-704

FIGURE 159

```
AGGGTCCCTTAGCCGGGCGCAGGGCGCGCAGCCCAGGCTGAGATCCGCGGCTTCCGTAGAAG
TGAGCATGGCTGGGCAGCGAGTGCTTCTTCTAGTGGGCTTCCTTCTCCCTGGGGTCCTGCTC
TCAGAGGCTGCCAAAATCCTGACAATATCTACAGTAGGTGGAAGCCATTATCTACTGATGGA
CCGGGTTTCTCAGATTCTTCAAGATCACGGTCATAATGTCACCATGCTTAACCACAAAAGAG
GTCCTTTTATGCCAGATTTTAAAAAGGAAGAAAAATCATATCAAGTTATCAGTTGGCTTGCA
CCTGAAGATCATCAAAGAGAATTTAAAAAGAGTTTTGATTTCTTTCTGGAAGAAACTTTAGG
TGGCAGAGGAAAATTTGAAAACTTATTAAATGTTCTAGAATACTTGGCGTTGCAGTGCAGTC
ATTTTTTAAATAGAAAGGATATCATGGATTCCTTAAAGAATGAGAACTTCGACATGGTGATA
GTTGAAACTTTTGACTACTGTCCTTTCCTGATTGCTGAGAAGCTTGGGAAGCCATTTGTGGC
CATTCTTTCCACTTCATTCGGCTCTTTGGAATTTGGGCTACCAATCCCCTTGTCTTATGTTC
CAGTATTCCGTTCCTTGCTGACTGATCACATGGACTTCTGGGGCCGAGTGAAGAATTTTCTG
ATGTTCTTTAGTTTCTGCAGGAGGCAACAGCACATGCAGTCTACATTTGACAACACCATCAA
GGAACATTTCACAGAAGGCTCTAGGCCAGTTTTGTCTCATCTTCTACTGAAAGCAGAGTTGT
GGTTCATTAACTCTGACTTTGCCTTTGATTTTGCTCGACCTCTGCTTCCCAACACTGTTTAT
GTTGGAGGCTTGATGGAAAAACCTATTAAACCAGTACCACAAGACTTGGAGAACTTCATTGC
CAAGTTTGGGGACTCTGGTTTTGTCCTTGTGACCTTGGGCTCCATGGTGAACACCTGTCAGA
ATCCGGAAATCTTCAAGGAGATGAACAATGCCTTTGCTCACCTACCCCAAGGGGTGATATGG
AAGTGTCAGTGTTCTCATTGGCCCAAAGATGTCCACCTGGCTGCAAATGTGAAAATTGTGGA
CTGGCTTCCTCAGAGTGACCTCCTGGCTCACCCAAGCATCCGTCTGTTTGTCACCCACGGCG
GGCAGAATAGCATAATGGAGGCCATCCAGCATGGTGTGCCCATGGTGGGGATCCCTCTCTTT
GGAGACCAGCCTGAAAACATGGTCCGAGTAGAAGCCAAAAGTTTGGTGTTTCTATTCAGTT
AAAGAAGCTCAAGGCAGAGACATTGGCTCTTAAGATGAAACAAATCATGGAAGACAAGAGAT
ACAAGTCCGCGGCAGTGGCTGCCAGTGTCATCCTGCGCTCCCACCCGCTCAGCCCCACACAG
CGGCTGGTGGGCTGGATTGACCACGTCCTCCAGACAGGGGGCGCGACGCACCTCAAGCCCTA
TGTCTTTCAGCAGCCCTGGCATGAGCAGTACCTGTTCGACGTTTTTGTGTTTCTGCTGGGGC
TCACTCTGGGGACTCTATGGCTTTGTGGGAAGCTGCTGGGCATGGCTGTCTGGTGGCTGCGT
GGGGCCAGAAAGGTGAAGGAGACATAAGGCCAGGTGCAGCCTTGGCGGGGTCTGTTTGGTGG
GCGATGTCACCATTTCTAGGGAGCTTCCCACTAGTTCTGGCAGCCCCATTCTCTAGTCCTTC
TAGTTATCTCCTGTTTTCTTGAAGAACAGGAAAAATGGCCAAAAATCATCCTTTCCACTTGC
TAATTTTGCTACAAATTCATCCTTACTAGCTCCTGCCTGCTAGCAGAAATCTTTCCAGTCCT
CTTGTCCTCCTTTGTTTGCCATCAGCAAGGGCTATGCTGTGATTCTGTCTCTGAGTGACTTG
GACCACTGACCCTCAGATTTCCAGCCTTAAAATCCACCTTCCTTCTCATGCGCCTCTCCGAA
TCACACCCTGACTCTTCCAGCCTCCATGTCCAGACCTAGTCAGCCTCTCTCACTCCTGCCCC
TACTATCTATCATGGAATAACATCCAAGAAAGACACCTTGCATATTCTTTCAGTTTCTGTTT
TGTTCTCCCACATATTCTCTTCAATGCTCAGGAAGCCTGCCCTGTGCTTGAGAGTTCAGGGC
CGGACACAGGCTCACAGGTCTCCACATTGGGTCCCTGTCTCTGGTGCCCACAGTGAGCTCCT
TCTTGGCTGAGCAGGCATGGAGACTGTAGGTTTCCAGATTTCCTGAAAAATAAAAGTTTACA
GCGTTATCTCTCCCCAACCTCACTAA
```

FIGURE 160

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71169

><subunit 1 of 1, 523 aa, 1 stop

><MW: 59581, pI: 8.68, NX(S/T): 1

MAGQRVLLLVGFLLPGVLLSEAAKILTISTVGGSHYLLMDRVSQILQDHGHNVTMLNHKRGP

FMPDFKKEEKSYQVISWLAPEDHQREFKKSFDFFLEETLGGRGKFENLLNVLEYLALQCSHF

LNRKDIMDSLKNENFDMVIVETFDYCPFLIAEKLGKPFVAILSTSFGSLEFGLPIPLSYVPV

FRSLLTDHMDFWGRVKNFLMFFSFCRRQQHMQSTFDNTIKEHFTEGSRPVLSHLLLKAELWF

INSDFAFDFARPLLPNTVYVGGLMEKPIKPVPQDLENFIAKFGDSGFVLVTLGSMVNTCQNP

EIFKEMNNAFAHLPQGVIWKCQCSHWPKDVHLAANVKIVDWLPQSDLLAHPSIRLFVTHGGQ

NSIMEAIQHGVPMVGIPLFGDQPENMVRVEAKKFGVSIQLKKLKAETLALKMKQIMEDKRYK

SAAVAASVILRSHPLSPTQRLVGWIDHVLQTGGATHLKPYVFQQPWHEQYLFDVFVFLLGLT

LGTLWLCGKLLGMAVWWLRGARKVKET

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 483-504

FIGURE 161

```
GGGCTGTTGATTTGTGGGGGATTTTGAAGAGAGGAGGAATAGGAGGAAGGGGTTGAGGGGCT
GCCTCTGGCATATGCACACACTCACACATTCTGTCACACCCGTCACACACACATACCATGTT
CTCCATCCCCCCAGGTCCAGCCCTCAGTGCTGTCCCATCCAGCAGGGCTACCCTGAAGCTCT
GGCTGCAGCCCTCCCGTCCAGTGGGCAGGCGGCTTCATCCCTCCTTTCTCTCCCAAAGCCCA
ACTGCTGTCACTGCATGCTCTGCCAAGGAGGAGGGAACTGCAGTGACAGCAGGAGTAAGAGT
GGGAGGCAGGACAGAGCTGGGACACAGGTATGGAGAGGGGGTTCAGCGAGCCTAGAGAGGGC
AGACTATCAGGGTGCCGGCGGTGAGAATCCAGGGAGAGGAGCGGAAACAGAAGAGGGGCAGA
AGACCGGGGCACTTGTGGGTTGCAGAGCCCCTCAGCCATGTTGGGAGCCAAGCCACACTGGC
TACCAGGTCCCCTACACAGTCCCGGGCTGCCCTTGGTTCTGGTGCTTCTGGCCCTGGGGGCC
GGGTGGGCCCAGGAGGGGGTCAGAGCCCGTCCTGCTGGAGGGGGAGTGCCTGGTGGTCTGTGA
GCCTGGCCGAGCTGCTGCAGGGGGGGCCCGGGGGAGCAGCCCTGGGAGAGGCACCCCCTGGGC
GAGTGGCATTTGCTGCGGTCCGAAGCCACCACCATGAGCCAGCAGGGGAAACCGGCAATGGC
ACCAGTGGGGCCATCTACTTCGACCAGGTCCTGGTGAACGAGGGCGGTGGCTTTGACCGGGC
CTCTGGCTCCTTCGTAGCCCCTGTCCGGGGTGTCTACAGCTTCCGGTTCCATGTGGTGAAGG
TGTACAACCGCCAAACTGTCCAGGTGAGCCTGATGCTGAACACGTGGCCTGTCATCTCAGCC
TTTGCCAATGATCCTGACGTGACCCGGGAGGCAGCCACCAGCTCTGTGCTACTGCCCTTGGA
CCCTGGGGACCGAGTGTCTCTGCGCCTGCGTCGGGGGAATCTACTGGGTGGTTGGAAATACT
CAAGTTTCTCTGGCTTCCTCATCTTCCCTCTCTGAGGACCCAAGTCTTTCAAGCACAAGAAT
CCAGCCCCTGACAACTTTCTTCTGCCCTCTCTTGCCCCAGAAACAGCAGAGGCAGGAGAGAG
ACTCCCTCTGGCTCCTATCCCACCTCTTTGCATGGGACCCTGTGCCAAACACCCAAGTTTAA
GAGAAGAGTAGAGCTGTGGCATCTCCAGACCAGGCCTTTCCACCCACCCACCCCCCAGTTACC
CTCCCAGCCACCTGCTGCATCTGTTCCTGCCTGCAGCCCTAGGATCAGGGCAAGGTTTGGCA
AGAAGGAAGATCTGCACTACTTTGCGGCCTCTGCTCCTCCGGTTCCCCCCACCCCAGCTTCCT
GCTCAATGCTGATCAGGGACAGGTGGCGCAGGTGAGCCTGACAGGCCCCCACAGGAGCCCAG
ATGGACAAGCCTCAGCGTACCCTGCAGGCTTCTTCCTGTGAGGAAAGCCAGCATCACGGATC
TCAGCCAGCACCGTCAGAAGCTGAGCCAGCACCGTATGGGCTAGGGTGGGAGGCTCAGCCAC
AGGCAGAAGGGTGGGAAGGGCCTGGAGTCTGTGGCTGGTGAGGAAGGAAGGAGGGTGTATTG
TCTAGACTGAACATGGTACACATTCTGCATGTATAGCAGAGCAGCCAGCAGGTAGCAATCCT
GGCTGTCCTTCTATGCTGGATCCCAGATGGACTCTGGCCCTTACCTCCCCACCTGAGATTAG
GGTGAGTGTGTTTGCTCTGGCTGAGAGCAGAGCTGAGAGCAGGTATACAGAGCTGGAAGTGG
ACCATGGAAAACATCGATAACCATGCATCCTCTTGCTTGGCCACCTCCTGAAACTGCTCCAC
CTTTGAAGTTTGAACTTTAGTCCCTCCACACTCTGACTGCTGCCTCCTTCCTCCCAGCTCTC
TCACTGAGTTATCTTCACTGTACCTGTTCCAGCATATCCCCACTATCTCTCTTTCTCCTGAT
CTGTGCTGTCTTATTCTCCTCCTTAGGCTTCCTATTACCTGGGATTCCATGATTCATTCCTT
CAGACCCTCTCCTGCCAGTATGCTAAACCCTCCCTCTCTCTTTCTTATCCCGCTGTCCCATT
GGCCCAGCCTGGATGAATCTATCAATAAAACAACTAGAGAATGGTGGTCAGTGAGACACTAT
AGAATTACTAAGGAGAAGATGCCTCTGGAGTTTGGATCGGGTGTTACAGGTACAAGTAGGTA
TGTTGCAGAGGAAAATAAATATCAAACTGTATACTAAAATTAAAAA
```

FIGURE 162

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71180

><subunit 1 of 1, 205 aa, 1 stop

><MW: 21521, pI: 7.07, NX(S/T): 1

MLGAKPHWLPGPLHSPGLPLVLVLLALGAGWAQEGSEPVLLEGECLVVCEPGRAAAGGPGGA

ALGEAPPGRVAFAAVRSHHHEPAGETGNGTSGAIYFDQVLVNEGGGFDRASGSFVAPVRGVY

SFRFHVVKVYNRQTVQVSLMLNTWPVISAFANDPDVTREAATSSVLLPLDPGDRVSLRLRRG

NLLGGWKYSSFSGFLIFPL

Signal peptide:

amino acids 1-32

164/270

FIGURE 163

GCTGTTTCTCTCGCGCCACCACTGGCCGCCGGCCGCAGCTCCAGGTGTCCTAGCCGCCCAGC

CTCGACGCCGTCCCGGGACCCCTGTGCTCTGCGCGAAGCCCTGGCCCCGGGGGCCGGGGCAT

GGGCCAGGGGCGCGGGGTGAAGCGGCTTCCCGCGGGGCCGTGACTGGGCGGGCTTCAGCC<u>AT</u>

<u>G</u>AAGACCCTCATAGCCGCCTACTCCGGGGTCCTGCGCGGCGAGCGTCAGGCCGAGGCTGACC

GGAGCCAGCGCTCTCACGGAGGACCTGCGCTGTCGCGCGAGGGGTCTGGGAGATGGGGCACT

GGATCCAGCATCCTCTCCGCCCTCCAGGACCTCTTCTCTGTCACCTGGCTCAATAGGTCCAA

GGTGGAAAAGCAGCTACAGGTCATCTCAGTGCTCCAGTGGGTCCTGTCCTTCCTTGTACTGG

GAGTGGCCTGCAGTGCCATCCTCATGTACATATTCTGCACTGATTGCTGGCTCATCGCTGTG

CTCTACTTCACTTGGCTGGTGTTTGACTGGAACACACCCAAGAAAGGTGGCAGGAGGTCACA

GTGGGTCCGAAACTGGGCTGTGTGGCGCTACTTTCGAGACTACTTTCCCATCCAGCTGGTGA

AGACACACAACCTGCTGACCACCAGGAACTATATCTTTGGATACCACCCCCATGGTATCATG

GGCCTGGGTGCCTTCTGCAACTTCAGCACAGAGGCCACAGAAGTGAGCAAGAAGTTCCCAGG

CATACGGCCTTACCTGGCTACACTGGCAGGCAACTTCCGAATGCCTGTGTTGAGGGAGTACC

TGATGTCTGGAGGTATCTGCCCTGTCAGCCGGGACACCATAGACTATTTGCTTTCAAAGAAT

GGGAGTGGCAATGCTATCATCATCGTGGTCGGGGGTGCGGCTGAGTCTCTGAGCTCCATGCC

TGGCAAGAATGCAGTCACCCTGCGGAACCGCAAGGGCTTTGTGAAACTGGCCCTGCGTCATG

GAGCTGACCTGGTTCCCATCTACTCCTTTGGAGAGAATGAAGTGTACAAGCAGGTGATCTTC

GAGGAGGGCTCCTGGGGCCGATGGGTCCAGAAGAAGTTCCAGAAATACATTGGTTTCGCCCC

ATGCATCTTCCATGGTCGAGGCCTCTTCTCCTCCGACACCTGGGGGCTGGTGCCCTACTCCA

AGCCCATCACCACTGTTGTGGGAGAGCCCATCACCATCCCCAAGCTGGAGCACCCAACCCAG

CAAGACATCGACCTGTACCACACCATGTACATGGAGGCCCTGGTGAAGCTCTTCGACAAGCA

CAAGACCAAGTTCGGCCTCCCGGAGACTGAGGTCCTGGAGGTGAAC<u>TGA</u>GCCAGCCTTCGGG

GCCAATTCCCTGGAGGAACCAGCTGCAAATCACTTTTTTGCTCTGTAAATTTGGAAGTGTCA

TGGGTGTCTGTGGGTTATTTAAAAGAAATTATAACAATTTTGCTAAACCAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAA

FIGURE 164

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71184
><subunit 1 of 1, 388 aa, 1 stop
><MW: 43831, pI: 9.64, NX(S/T): 3
MKTLIAAYSGVLRGERQAEADRSQRSHGGPALSREGSGRWGTGSSILSALQDLFSVTWLNRS
KVEKQLQVISVLQWVLSFLVLGVACSAILMYIFCTDCWLIAVLYFTWLVFDWNTPKKGGRRS
QWVRNWAVWRYFRDYFPIQLVKTHNLLTTRNYIFGYHPHGIMGLGAFCNFSTEATEVSKKFP
GIRPYLATLAGNFRMPVLREYLMSGGICPVSRDTIDYLLSKNGSGNAIIIVVGGAAESLSSM
PGKNAVTLRNRKGFVKLALRHGADLVPIYSFGENEVYKQVIFEEGSWGRWVQKKFQKYIGFA
PCIFHGRGLFSSDTWGLVPYSKPITTVVGEPITIPKLEHPTQQDIDLYHTMYMEALVKLFDK
HKTKFGLPETEVLEVN
```

Important features of the protein:
Transmembrane domain:
amino acids 76-97

N-glycosylation sites.
amino acids 60-63, 173-176, 228-231

N-myristoylation sites.
amino acids 10-15, 41-46, 84-89, 120-125, 169-174, 229-234, 240-
245, 318-323, 378-383

FIGURE 165

```
GGGCGGCGGGATGGGGGCCGGGGGCGGCGGGCGCCGCACTCGCTGAGGCCCCGACGCAGGGCCGGGCCGGGCCCA
GGGCCGAGGAGCGCGGCGGCCAGAGCGGGGCCGCGGAGGCGACGCCGGGGACGCCCGCGCGACGAGCAGGTGGCG
GCGGCTGCAGGCTTGTCCAGCCGGAAGCCCTGAGGGCAGCTGTTCCCACTGGCTCTGCTGACCTTGTGCCTTGGA
CGGCTGTCCTCAGCGAGGGGCCGTGCACCCGCTCCTGAGCAGCGCCATGGGCCTGCTGGCCTTCCTGAAGACCCA
GTTCGTGCTGCACCTGCTGGTCGGCTTTGTCTTCGTGGTGAGTGGTCTGGTCATCAACTTCGTCCAGCTGTGCAC
GCTGGCGCTCTGGCCGGTCAGCAAGCAGCTCTACCGCCGCCTCAACTGCCGCCTCGCCTACTCACTCTGGAGCCA
ACTGGTCATGCTGCTGGAGTGGTGGTCCTGCACGGAGTGTACACTGTTCACGGACCAGGCCACGGTAGAGCGCTT
TGGGAAGGAGCACGCAGTCATCATCCTCAACCACAACTTCGAGATCGACTTCCTCTGTGGGTGGACCATGTGTGA
GCGCTTCGGAGTGCTGGGGAGCTCCAAGGTCCTCGCTAAGAAGGAGCTGCTCTACGTGCCCCTCATCGGCTGGAC
GTGGTACTTTCTGGAGATTGTGTTCTGCAAGCGGAAGTGGGAGGAGGACCGGGACACCGTGGTCGAAGGGCTGAG
GCGCCTGTCGGACTACCCCGAGTACATGTGGTTTCTCCTGTACTGCGAGGGGACGCGCTTCACGGAGACCAAGCA
CCCGCGTTAGCATGGAGGTGGCGGCTGCTAAGGGGCTTCCTGTCCTCAAGTACCACCTGCTGCCGCGGACCAAGGG
CTTCACCACCGCAGTCAAGTGCCTCCGGGGGACAGTCGCAGCTGTCTATGATGTAACCCTGAACTTCAGAGGAAA
CAAGAACCCGTCCCTGCTGGGGATCCTCTACGGGAAGAAGTACGAGGCGGACATGTGCGTGAGGAGATTTCCTCT
GGAAGACATCCCGCTGGATGAAAAGGAAGCAGCTCAGTGGCTTCATAAACTGTACCAGGAGAAGGACGCGCTCCA
GGAGATATATAATCAGAAGGGCATGTTTCCAGGGGAGCAGTTTAAGCCTGCCCGGAGGCCGTGGACCCTCCTGAA
CTTCCTGTCCTGGGCCACCATTCTCCTGTCTCCCCTCTTCAGTTTTGTCTTGGGCGTCTTTGCCAGCGGATCACC
TCTCCTGATCCTGACTTTCTTGGGGTTTGTGGGAGCAGCTTCCTTTGGAGTTCGCAGACTGATAGGAGAATCGCT
TGAACCTGGGAGGTGGAGATTGCAGTGAGCTGAGATGGCATCACTGTACTCCAGCCTAGGCAACAGAGCAAGACT
CAGTCTCAAAAAAAAAAAAAAAACAAAAAAACCCCAGAAATTCTGGAGTTGAACTGTGTAGTTACTGACATGAAAA
ATTCACTAGAGGCTGAACAGCAGATTTGAGCAGGCAGAAAAAAAATCAGCAAGCTTGAAGATGGTACCTTGAGATT
TTTCAGGCTAATGAAAAAAGAATGAAGGAAAATTAACAGCCTCAGAGACCCATGGTGCACCGTCACACAAATCAA
CATATGCATGATGAGAGTCCCAGAAGGAGAGGAGAGAAAGGGTCAGAAAGAATGGCCACAAGCTGATGAAAAACA
GTAACCTACCCACTCAGGAAGCTCAGTGAACTCCAATGAGGATGAATATCAGAGATCCACACCTAGATATTTCAT
AATCAAAGTGTCAAATGACAAAGAATCTTGAAAGCAGCAAGAGATGAGCAACTTATCTTGTTCAAAGGATCTTTG
ATCAGATTAACAGCTCATTTCTCCTCAGAAATCATGGGAGCCAGGAGATAGTGGGATGAACACTGTTGAAGGCAA
AACCTTCAACTGTAATTATTGGACTTTTGAGTCTTAGATGGTCCTGACCTCTTTGTCTTCAGGGACAGTTTTTCA
ATTTAATCCCTAATAACAATTAGTCAAGCTTCCTTGACCTGTAGGAAGGCCTGTCTTTAGGCCGGGCACAGTGGC
TTACACCTGTAATCCCAGCACTTTGGGAGGCCCAGACGGGTGGATCATTTGGGGTCAGGCTGATCTCAAACTCCT
GAGTTCAGGTGATCTGCCCGCCTCAGCCTCCCAAAGTGTTGTGATTGCAGGCGTGAGCCACTGCGCCTGGCCGGA
ATTTCTTTTTAAGGCTGAATGATGGGGGCCAGGCACGATGGCTCACGCCTGTGATCCCAAGTAGCTTGGATTGTA
AACATGCACCACCATGCCTGGCTAATTTTTGTATTTTTAGTAGAGACGTGTTAGCCAGGCTGGTCTCGATCTCCT
GACCTCAAGTGACCACCTGCCTCAGCCTCCCAAAGTACTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCTTGA
GCATCTTGTGATGTGCTTATTGGCCATTTGTATATCTTCTATCTTCTTTGGGGAAATGTCTGTTCAAGTCCTTTG
CCTTTTTAAATTTTTATTATTTATTTATTTATTTATTTTGAGACAGGGTCTTGTTCTGTTGCCCAGGCTGGAGTA
CAGTGGCACAGTCTTGGCTCACTGCAGCCTCGACCTCCTGGGCTGCAGTGATCCTCCCACCTCAGCCTCCCTTGT
AGCTGTATTTTTTTGTATTTTGTATTTTGTAGCTGTAGTTTTTGTATTTTTTGTGGAGACAGCATTTCACCATGA
TGCCCAGGCTGGTCTTGAACTCCTGAGCTCAAGTGATCTGCCTGCTTCAGCCTCCCAAAGTGCTGGGATTACAGA
CATGAGCCACTGCACCTGGCAAACTCCCAAAATTCAACACACACACACAAAAAACCACCTGATTCAAAATGGGCA
GAGGGGCCGGGTGTGGCCCCAACTACCAGGGAGACTGAAGTGGGAGGATCGCTTGGGCATGAGAAGTCGAGGCTG
CAGTGAGTCGAGGTTGTGCGACTGCATTCCAGCCTGGACAACAGAGTGAGACCCTGTCTC
```

FIGURE 166

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71213

><subunit 1 of 1, 368 aa, 1 stop

><MW: 42550, pI: 9.11, NX(S/T): 1

MGLLAFLKTQFVLHLLVGFVFVVSGLVINFVQLCTLALWPVSKQLYRRLNCRLAYSLWSQLV

MLLEWWSCTECTLFTDQATVERFGKEHAVIILNHNFEIDFLCGWTMCERFGVLGSSKVLAKK

ELLYVPLIGWTWYFLEIVFCKRKWEEDRDTVVEGLRRLSDYPEYMWFLLYCEGTRFTETKHR

VSMEVAAAKGLPVLKYHLLPRTKGFTTAVKCLRGTVAAVYDVTLNFRGNKNPSLLGILYGKK

YEADMCVRRFPLEDIPLDEKEAAQWLHKLYQEKDALQEIYNQKGMFPGEQFKPARRPWTLLN

FLSWATILLSPLFSFVLGVFASGSPLLILTFLGFVGAASFGVRRLIGESLEPGRWRLQ

Important features of the protein:

Signal peptide:

amino acids 1-25

Transmembrane domains:

amino acids 307-323, 335-352

Tyrosine kinase phosphorylation sites.

amino acids 160-168, 161-169

FIGURE 167

GATATTCTTTATTTTTAAGAATCTGAAGTACT**ATG**CATCACTCCCTCCAATGTCCTGGGGCA

GCCACCAGGCATATTCATCTTTGTGTGTGTTTTTCTTTTGCTTTAGCACTGGGGCACTTCTT

GCTTATTTCTTTGGTAGGAAAGGGGCTCAGTTTGTCTTGTGGGGTTGGTGGCAGGCAGGCCG

GCTTACGCCTGATACGGCCCTGGGTTAGAAGGGAAGGGAAGATAAACTTTTATACAAATGGG

GATAGCTGGGGTCTGAGACCTGCTTCCTCAGTAAAATTCCTGGGATCTGCCTATACCTTCTT

TTCTCTAACCTGGCATACCCTGCTTAAAGCCTCTCAGGGCTTCTCTCTGTTCTTAGGATCAA

AGTATTTAGAGCTACAAGAGCCCTCATGGTCTGGCCCCTGCCCCCCTGGCCAGCTTCATTGT

ACATGTGGTGTTCTCTTGTCGTTCCTG**TAA**TGTGGTATGCCATGGGGTCTTTGCACAAGCCT

TTCCTCTTTGGCTGGACACTGTTCCCTGCCCCCCCCATACTCTTCCTACTTAATATGTAGTC

ATCCTGCAGATTTCAATTCTAACATCATTTTCTCCAGGGATCCTGGCCTGACAGAATCTCAT

CTTGTTTAATGCTCTCATAAGACCACTTGTTTCCCTTTTGCAGCACTTGCCACTCAGTTGTA

TCTTTATGTGCGTTTGTGGTTGTATGGGTTGTGTCTGTTCCCCAGAATGCCCAGCTCTGAGC

TGCGTGAGGGTCAAGGGCATTGCTGTGCCTGCCAGGTATAGTGCCTACATGTGGTGGGTGCT

CATGTTTTAGAGACTAAATGGAGGAGGAGATGAGGAAAAGATTGAAATCTCTCAGTTCACCA

GATGGTGTAGGGCCCAGCATTGTAAATTCACACGTTGACTGTGCTTGTGAATTATCTGGGGA

TGCAGGTCCTGATTCAGTAGGCCCAGGTTGGGCATCTCTAACAAACTCCCACGTGATGCTGA

TGCTGGTCCTATGAACTATACTAAATAGTAAGAATCTATGGAGCCAGGCTGGGCATGGTGGC

TCACACCTATGATCCCAGCACTTTGGGAGGCTGAGGCAGGCTGATCACCTGGAGTCAGGATT

TCAAGACTAGCCTGGCCAACATGGTGGAACCCCATCTGTACTAAAAATACACAAATTAGCTG

GGCATGGTGGCACATGCCTGTAGTCCCAGCTACTTGGGAGGCTGAAGCAAGAGAATCGCTTG

AACCTGGGAGGCGGAGGTTGCAGTGAGCCGAGATCAGGCCACTGTATTCCAACCAGGGTGAC

AGAGTGAGACTCTATGTCCAAAAAAAAAAAAAA

FIGURE 168

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71234

><subunit 1 of 1, 143 aa, 1 stop

><MW: 15624, pI: 9.58, NX(S/T): 0

MHHSLQCPGAATRHIHLCVCFSFALALGHFLLISLVGKGLSLSCGVGGRQAGLRLIRPWVRR
EGKINFYTNGDSWGLRPASSVKFLGSAYTFFSLTWHTLLKASQGFSLFLGSKYLELQEPSWS
GPCPPGQLHCTCGVLLSFL

Important features of the protein:

Signal peptide:

amino acids 1-28

FIGURE 169

```
GGCTGGACTGGAACTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGA
TTATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTCAGCAACTAAAAAAGCCAC
AGGAGTTGAACTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTAC
ATTAAAATCTGTTTTTTGTTCTCTTGTAACTAGCCTTTACCTTCCTAACACAGAGGATCTGT
CACTGTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCAC
ACCGTCCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTC
ACCAACTGTCTCACGTCTGGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTA
GCTGCGGCTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTT
CATAGGCGATGGCTCCCACTGCCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGG
CCAGGACGGGCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCGCCCAT
CTAACCTTTTCATGTCCTGCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCCAAGG
AACCCAGAGCTTGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTT
TATGTGACAGGACTTGCATTCTCCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCA
GGGAAGGAACTTGTGCCAAATTATGGGTCAGAAAAGATGGAGGTGTTGGGTTATCACAAGGC
ATCGAGTCTCCTGCATTCAGTGGACATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACT
CGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGATCCACGTACCAGCTGCTG
AAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACCAG
CCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCTCCTTCCCTC
TGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGCTAATG
GCTCAGTGTTGGCCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCG
AACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCT
CAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGT
TCAAATGATCTCCAAGGGCCCTTATACCCCAGGAGACTTTGATTTGAATTTGAAACCCCAAA
TCCAAACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATG
CCAACATTTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTG
GCCAACATGGTGAAACCCCTGTCTCTACTAAAAATACAAAAAAACTAGCCAGGCATGGTGGT
GTGTGCCTGTATCCCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGT
GAAGGAGGCTGAGACAGGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGA
AAAAATAAAAAAAGAATTATGGTTATTTGTAA
```

FIGURE 170

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71277

><subunit 1 of 1, 109 aa, 1 stop

><MW: 11822, pI: 8.63, NX(S/T): 0

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTA
SPCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:
amino acids 1-15

FIGURE 171

GCGGGCCCGCGAGTCCGAGACCTGTCCCAGGAGCTCCAGCTCACGTGACCTGTCACTGCCTC

CCGCCGCCTCCTGCCCGCGCC**ATG**ACCCAGCCGGTGCCCCGGCTCTCCGTGCCCGCCGCGCT

GGCCCTGGGCTCAGCCGCACTGGGCGCCGCCTTCGCCACTGGCCTCTTCCTGGGGAGGCGGT

GCCCCCCATGGCGAGGCCGGCGAGAGCAGTGCCTGCTTCCCCCCGAGGACAGCCGCCTGTGG

CAGTATCTTCTGAGCCGCTCCATGCGGGAGCACCCGGCGCTGCGAAGCCTGAGGCTGCTGAC

CCTGGAGCAGCCGCAGGGGGATTCTATGATGACCTGCGAGCAGGCCCAGCTCTTGGCCAACC

TGGCGCGGCTCATCCAGGCCAAGAAGGCGCTGGACCTGGGCACCTTCACGGGCTACTCCGCC

CTGGCCCTGGCCCTGGCGCTGCCCGCGGACGGGCGCGTGGTGACCTGCGAGGTGGACGCGCA

GCCCCCGGAGCTGGGACGGCCCCTGTGGAGGCAGGCCGAGGCGGAGCACAAGATCGACCTCC

GGCTGAAGCCCGCCTTGGAGACCCTGGACGAGCTGCTGGCGGCGGGCGAGGCCGGCACCTTC

GACGTGGCCGTGGTGGATGCGGACAAGGAGAACTGCTCCGCCTACTACGAGCGCTGCCTGCA

GCTGCTGCGACCCGGAGGCATCCTCGCCGTCCTCAGAGTCCTGTGGCGCGGGAAGGTGCTGC

AACCTCCGAAAGGGGACGTGGCGGCCGAGTGTGTGCGAAACCTAAACGAACGCATCCGGCGG

GACGTCAGGGTCTACATCAGCCTCCTGCCCCTGGGCGATGGACTCACCTTGGCCTTCAAGAT

C**TAG**GGCTGGCCCCTAGTGAGTGGGCTCGAGGGAGGGTTGCCTGGGAACCCCAGGAATTGAC

CCTGAGTTTTAAATTCGAAAATAAAGTGGGGCTGGGACACAAAAAAAAAAAAAAAAAAAA

FIGURE 172

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71282

><subunit 1 of 1, 262 aa, 1 stop

><MW: 28809, pI: 8.80, NX(S/T): 1

MTQPVPRLSVPAALALGSAALGAAFATGLFLGRRCPPWRGRREQCLLPPEDSRLWQYLLSRS

MREHPALRSLRLLTLEQPQGDSMMTCEQAQLLANLARLIQAKKALDLGTFTGYSALALALAL

PADGRVVTCEVDAQPPELGRPLWRQAEAEHKIDLRLKPALETLDELLAAGEAGTFDVAVVDA

DKENCSAYYERCLQLLRPGGILAVLRVLWRGKVLQPPKGDVAAECVRNLNERIRRDVRVYIS

LLPLGDGLTLAFKI

Important features of the protein:

Signal peptide:

amino acids 1-25

Transmembrane domains:

amino acids 8-30, 109-130

N-glycosylation site.

amino acids 190-193

Tyrosine kinase phosphorylation site.

amino acids 238-246

N-myristoylation sites.

amino acids 22-27, 28-33, 110-115, 205-210, 255-260

Amidation sites.

amino acids 31-34, 39-42

FIGURE 173

```
CCGCCGCCGCAGCCGCTACCGCCGCTGCAGCCGCTTTCCGCGGCCTGGGCCTCTCGCCGTCA
GCATGCCACACGCCTTCAAGCCCGGGGACTTGGTGTTCGCTAAGATGAAGGGCTACCCTCAC
TGGCCTGCCAGGATCGACGACATCGCGGATGGCGCCGTGAAGCCCCCACCCAACAAGTACCC
CATCTTTTTCTTTGGCACACACGAAACAGCCTTCCTGGGACCCAAGGACCTGTTCCCCTACG
ACAAATGTAAAGACAAGTACGGGAAGCCCAACAAGAGGAAAGGCTTCAATGAAGGGCTGTGG
GAGATCCAGAACAACCCCCACGCCAGCTACAGCGCCCCTCCGCCAGTGAGCTCCTCCGACAG
CGAGGCCCCGAGGCCAACCCCGCCGACGGCAGTGACGCTGACGAGGACGATGAGGACCGGG
GGGTCATGGCCGTCACAGCGGTAACCGCCACAGCTGCCAGCGACAGGATGGAGAGCGACTCA
GACTCAGACAAGAGTAGCGACAACAGTGGCCTGAAGAGGAAGACGCCTGCGCTAAAGATGTC
GGTCTCGAAACGAGCCCGAAAGGCCTCCAGCGACCTGGATCAGGCCAGCGTGTCCCCATCCG
AAGAGGAGAACTCGGAAAGCTCATCTGAGTCGGAGAAGACCAGCGACCAGGACTTCACACCT
GAGAAGAAAGCAGCGGTCCGGGCGCCACGGAGGGGCCCTCTGGGGGGACGGAAAAAAAGAA
GGCGCCGTCAGCCTCCGACTCCGACTCCAAGGCCGATTCGGACGGGGCCAAGCCTGAGCCGG
TGGCCATGGCGCGGTCGGCGTCCTCCTCCTCCTCTTCCTCCTCCTCCGACTCCGATGTG
TCTGTGAAGAAGCCTCCGAGGGGCAGGAAGCCAGCGGAGAAGCCTCTCCCGAAGCCGCGAGG
GCGGAAACCGAAGCCTGAACGGCCTCCGTCCAGCTCCAGCAGTGACAGTGACAGCGACGAGG
TGGACCGCATCAGTGAGTGGAAGCGGCGGGACGAGGCGCGGAGGCGCGAGCTGGAGGCCCGG
CGGCGGCGAGAGCAGGAGGAGGAGCTGCGGCGCCTGCGGGGAGCAGGAGAAGGAGGAGAAGGA
GCGGAGGCGCGAGCGGGCCGACCGCGGGGAGGCTGAGCGGGGCAGCGGCGGCAGCAGCGGGG
ACGAGCTCAGGGAGGACGATGAGCCCGTCAAGAAGCGGGGACGCAAGGGCCGGGGCCGGGGT
CCCCCGTCCTCCTCTGACTCCGAGCCCGAGGCCGAGCTGGAGAGAGAGGCCAAGAAATCAGC
GAAGAAGCCGCAGTCCTCAAGCACAGAGCCCGCCAGGAAACCTGGCCAGAAGGAGAAGAGAG
TGCGGCCCGAGGAGAAGCAACAAGCCAAGCCCGTGAAGGTGGAGCGGACCCGGAAGCGGTCC
GAGGGCTTCTCGATGGACAGGAAGGTAGAGAAGAAGAAAGAGCCCTCCGTGGAGGAGAAGCT
GCAGAAGCTGCACAGTGAGATCAAGTTTGCCCTAAAGGTCGACAGCCCGGACGTGAAGAGGT
GCCTGAATGCCCTAGAGGAGCTGGGAACCCTGCAGGTGACCTCTCAGATCCTCCAGAAGAAC
ACAGACGTGGTGGCCACCTTGAAGAAGATTCGCCGTTACAAAGCGAACAAGGACGTAATGGA
GAAGGCAGCAGAAGTCTATACCCCGGCTCAAGTCGCGGGTCCTCGGCCCAAAGATCGAGGCGG
TGCAGAAAGTGAACAAGGCTGGGATGGAGAAGGAGAAGGCCGAGGAGAAGCTGGCCGGGGGAG
GAGCTGGCCGGGGAGGAGGCCCCCCAGGAGAAGGCGGAGGACAAGCCCAGCACCGATCTCTC
AGCCCCAGTGAATGGCGAGGCCACATCACAGAAGGGGGAGAGCGCAGAGGACAAGGAGCACG
AGGAGGGTCGGGACTCGGAGGAGGGGCCAAGGTGTGGCTCCTCTGAAGACCTGCACGACAGC
GTACGGGAGGGTCCCGACCTGGACAGGCCTGGGAGCGACCGGCAGGAGCGCGAGAGGGCACG
GGGGGACTCGGAGGCCCTGGACGAGGAGAGCTGAGCCGCGGGCAGCCAGGCCCAGCCCCCGC
CCGAGCTCAGGCTGCCCCTCTCCTTCCCCGGCTCGCAGGAGAGCAGAGCAGAGAACTGTGGG
GAACGCTGTGCTGTTTGTATTTGTTCCCTTGGGTTTTTTTTTTCCTGCCTAATTTCTGTGATT
TCCAACCAACATGAAATGACTATAAACGGTTTTTTAATGA
```

FIGURE 174

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71286
><subunit 1 of 1, 671 aa, 1 stop
><MW: 74317, pI: 7.61, NX(S/T): 0
MPHAFKPGDLVFAKMKGYPHWPARIDDIADGAVKPPPNKYPIFFFGTHETAFLGPKDLFPYD
KCKDKYGKPNKRKGFNEGLWEIQNNPHASYSAPPPVSSSDSEAPEANPADGSDADEDDEDRG
VMAVTAVTATAASDRMESDSDSDKSSDNSGLKRKTPALKMSVSKRARKASSDLDQASVSPSE
EENSESSSESEKTSDQDFTPEKKAAVRAPRRGPLGGRKKKKAPSASDSDSKADSDGAKPEPV
AMARSASSSSSSSSSDSDVSVKKPPRGRKPAEKPLPKPRGRKPKPERPPSSSSSDSDSDEV
DRISEWKRRDEARRRELEARRRREQEEELRRLREQEKEEKERRRERADRGEAERGSGGSSGD
ELREDDEPVKKRGRKGRGRGPPSSSDSEPEAELEREAKKSAKKPQSSSTEPARKPGQKEKRV
RPEEKQQAKPVKVERTRKRSEGFSMDRKVEKKKEPSVEEKLQKLHSEIKFALKVDSPDVKRC
LNALEELGTLQVTSQILQKNTDVVATLKKIRRYKANKDVMEKAAEVYTRLKSRVLGPKIEAV
QKVNKAGMEKEKAEEKLAGEELAGEEAPQEKAEDKPSTDLSAPVNGEATSQKGESAEDKEHE
EGRDSEEGPRCGSSEDLHDSVREGPDLDRPGSDRQERERARGDSEALDEES
```

Signal peptide:
amino acids 1-13

FIGURE 175

```
GTTGGTTCTCCTGGATCTTCACCTTACCAACTGCAGATCTTGGGACTCATCAGCCTCAATAATTATATTAAATTA
ACACCATTTGAAAGAGAACATTGTTTTCATCATGAATGCTAATAAAGATGAAAGACTTAAAGCCAGAAGCCAAGA
TTTTCACCTTTTTCCTGCTTTGATGATGCTAAGCATGACCATGTTGTTTCTTCCAGTCACTGGCACTTTGAAGCA
AAATATTCCAAGACTCAAGCTAACCTACAAAGACTTGCTGCTTTCAAATAGCTGTATTCCCTTTTTGGGTTCATC
AGAAGGACTGGATTTTCAAACTCTTCTCCTTAGATGAGGAAAGAGGCAGGCTGCTCTTGGGAGCCAAAGACCACAT
CTTTCTACTCAGTCTGGTTGACTTAAACAAAAATTTTAAGAAGATTTATTGGCCTGCTGCAAAGGAACGGGTGGA
ATTATGTAAATTAGCTGGGAAAGATGCCAATACAGAATGTGCAAATTTCATCAGAGTACTTCAGCCCTATAACAA
AACTCACATATATGTGTGTGGAACTGGAGCATTTCATCCAATATGTGGGTATATTGATCTTGGAGTCTACAAGGA
GGATATTATATTCAAACTAGACACACATAATTTGGAGTCTGGCAGACTGAAATGTCCTTTCGATCCTCAGCAGCC
TTTTGCTTCAGTAATGACAGATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCCTTGGCAAAGATACTGCATT
CACTCGATCCCTTGGGCCTACTCATGACCACCACTACATCAGAACTGACATTTCAGAGCACTACTGGCTCAATGG
AGCAAAATTTATTGGAACTTTCTTCATACCAGACACCTACAATCCAGATGATGATAAAATATATTTCTTCTTTCG
TGAATCATCTCAAGAAGGCAGTACCTCCGATAAAACCATCCTTTCTCGAGTTGGAAGAGTTTGTAAGAATGATGT
AGGAGGACAACGCAGCCTGATAAACAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTCCTGGAAG
TGATGGGGCAGATACTTACTTTGATGAGCTTCAAGATATTTATTTACTCCCCACAAGAGATGAAAGAAATCCTGT
AGTATATGGAGTCTTTACTACAACCAGCTCCATCTTCAAAGGCTCTGCTGTTTGTGTGTATAGCATGGCTGACAT
CAGAGCAGTTTTTAATGGTCCATATGCTCATAAGGAAAGTGCAGACCATCGTTGGGTGCAGTATGATGGGAGAAT
TCCTTATCCACGGCCTGGTACATGTCCAAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGA
TGATGTCATCAGTTTCATAAAGCGGCACTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGAGGACCAACGTT
CAAGAGAATCAATGTGGATTACAGAGCTGACACAGATAGTGGTGGATCATGTCATTGCAGAAGATGGCCAGTACGA
TGTAATGTTTCTTGGAACAGACATTGGAACTGTCCTCAAAGTTGTCAGCATTTCAAAGGAAAAGTGGAATATGGA
AGAGGTAGTGCTGGAGGAGTTGCAGATATTCAAGCACTCATCAATCATCTTGAACATGGAATTGTCTCTGAAGCA
GCAACAATTGTACATTGGTTCCCGAGATGGATTAGTTCAGCTCTCCTTGCACAGATGCGACACTTATGGGAAAGC
TTGCGCAGACTGTTGTCTTGCCAGAGACCCCTACTGTGCCTGGGATGGAAATGCATGCTCTCGATATGCTCCTAC
TTCTAAAAGGAGAGCTAGACGCCAAGATGTAAAATATGGCGACCCAATCACCCAGTGCTGGGACATCGAAGACAG
CATTAGTCATGAAACTGCTGATGAAAAGGTGATTTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATACC
TAAATCCCAACAAGCAACTATTAAATGGTATATCCAGAGGTCAGGGGATGAGCATCGAGAGGAGTTGAAGCCCGA
TGAAAGAATCATCAAAACGGAATATGGGCTACTGATTCGAAGTTTGCAGAAGAAGGATTCTGGGATGTATTACTG
CAAAGCCCAGGAGCACACTTTCATCCACACCATAGTGAAGCTGACTTTGAATGTCATTGAGAATGAACAGATGGA
AAATACCCAGAGGGCGAGCATGAGGAGGGGCAGGTCAAGGATCTATTGGCTGAGTCACGGTTGAGATACAAAGA
CTACATCCAAATCCTTAGCAGCCCAAACTTCAGCCTCGACCAGTACTGCGAACAGATGTGGCACAGGGAGAAGCG
GAGACAGAGAAACAAGGGGGGCCCAAAGTGGAAGCACATGCAGGAAATGAAGAAGAAACGAAATCGAAGACATCA
CAGAGACCTGGATGAGCTCCCTAGAGCTGTAGCCACGTAGTTTTCTACTTAATTTAAAGAAAAGAATTCCTTACC
TATAAAAACATTGCCTTCTGTTTTGTATATCCCTTATAGTAATTCATAAATGCTTCCCATGGAGTTTTGCTAAGG
CACAAGACAATAATCTGAATAAGACAATATGTGATGAATATAAGAAAGGGCAAAAAATTCATTTGAACCAGTTTT
CCAAGAACAAATCTTGCACAAGCAAAGTATAAGAATTATCCTAAAAATAGGGGGTTTACAGTTGTAAATGTTTTA
TGTTTTGAGTTTTGGAATTTATTGTCATGTAAATAGTTGAGCTAAGCAAGCCCCGAATTTGATAGTGTATAAGGT
GCTTTATTCCCTCGAATGTCCATTAAGCATGGAATTTACCATGCAGTTGTGCTATGTTCTTATGAACAGATATAT
CATTCCTATTGAGAACCAGCTACCTTGTGGTAGGGAATAAGAGGTCAGACACACAAATTAAGACAACTCCCATTATC
AACAGGAACTTTCTCAGTGAGCCATTCACTCCTGGAGAATGGTATAGGAATTTGGAGAGGTGCATTATTTCTTTC
TGGCCACTGGGGTTAAATTTAGTGTACTACAACATTGATTTACTGAAGGGCACTAATGTTTCCCCCAGGATTTCT
ATTGACTAGTCAGGAGTAACAGGTTCACAGAGAGAAGTTGGTGCTTAGTTATGTGTTTTTTAGAGTATATACTAA
GCTCTACAGGGACAGAATGCTTAATAAATACTTTAATAAGATATGGGAAAATATTTTAATAAAACAAGGAAAACA
TAATGATGTATAATGCATCCTGATGGGAAGGCATGCAGATGGGATTTGTTAGAAGACAGAAGGAAAGACAGCCAT
AAATTCTGGCTTTGGGGAAAACTCATATCCCCATGAAAAGGAAGAACAATCACAAATAAAGTGAGAGTAATGTAA
TGGAGCTCTTTTCACTAGGGTATAAGTAGCTGCCAATTTGTAATTCATCTGTTAAAAAAAATCTAGATTATAACA
AACTGCTAGCAAAATCTGAGGAAACATAAATTCTTCTGAAGAATCATAGGAAGAGTAGACATTTTATTTATAACC
AATGATATTTCAGTATATATTTTCTCTCTTTTAAAAAAATATTTATCATACTCTGTATATTATTTCTTTTTACTGC
CTTTATTCTCTCCTGTATATTGGATTTTGTGATTATATTTGAGTGAATAGGAGAAAACAATATATAACACACAGA
GAATTAAGAAAATGACATTTCTGGGGAGTGGGGATATATATTTGTTGAATAACAGAACGAGTGTAAAATTTTAAC
AACGGAAAGGGTTAAATTAACTCTTTGACATCTTCACTCAACCTTTTCTCATTGCTGAGTTAATCTGTTGTAATT
GTAGTATTGTTTTTGTAATTTAACAATAAATAAGCCTGCTACATGT
```

FIGURE 176

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71883

><subunit 1 of 1; 777 aa, 1 stop

><MW: 89651, pI: 7.97, NX(S/T): 3

MNANKDERLKARSQDFHLFPALMMLSMTMLFLPVTGTLKQNIPRLKLTYKDLLLSNSCIPFL

GSSEGLDFQTLLLLDEERGRLLLGAKDHIFLLSLVDLNKNFKKIYWPAAKERVELCKLAGKDA

NTECANFIRVLQPYNKTHIYVCGTGAFHPICGYIDLGVYKEDIIFKLDTHNLESGRLKCPFD

PQQPFASVMTDEYLYSGTASDFLGKDTAFTRSLGPTHDHHYIRTDISEHYWLNGAKFIGTFF

IPDTYNPDDDKIYFFFRESSQEGSTSDKTILSRVGRVCKNDVGGQRSLINKWTTFLKARLIC

SIPGSDGADTYFDELQDIYLLPTRDERNPVVYGVFTTTSSIFKGSAVCVYSMADIRAVFNGP

YAHKESADHRWVQYDGRIPYPRPGTCPSKTYDPLIKSTRDFPDDVISFIKRHSVMYKSVYPV

AGGPTFKRINVDYRLTQIVVDHVIAEDGQYDVMFLGTDIGTVLKVVSISKEKWNMEEVVLEE

LQIFKHSSIILNMELSLKQQQLYIGSRDGLVQLSLHRCDTYGKACADCCLARDPYCAWDGNA

CSRYAPTSKRRARRQDVKYGDPITQCWDIEDSISHETADEKVIFGIEFNSTFLECIPKSQQA

TIKWYIQRSGDEHREELKPDERIIKTEYGLLIRSLQKKDSGMYYCKAQEHTFIHTIVKLTLN

VIENEQMENTQRAEHEEGQVKDLLAESRLRYKDYIQILSSPNFSLDQYCEQMWHREKRRQRN

KGGPKWKHMQEMKKKRNRRHHRDLDELPRAVAT

Important features of the protein:
Signal peptide:
amino acids 1-36

N-glycosylation sites.
amino acids 139-142, 607-610, 724-727

Tyrosine kinase phosphorylation site.
amino acids 571-576

Gram-positive cocci surface proteins 'anchoring' hexapeptide.
amino acids 32-37

FIGURE 177

```
CCCTGACCTCCCTGAGCCACACTGAGCTGGAAGCCGCAGAGGTCATCCTGGAGCATGCCCACCGCGGGGAGCAGA
CAACCTCCCAGGTAAGCTGGGAGCAAGACCTGAAGCTGTTTCTTCAGGAGCCTGGTGTATTTTCCCCCACCCCAC
CTCAGCAGTTTCAGCCAGCAGGGACTGATCAGGTGTGTGTCCTGGAGTGGGGAGCAGAAGGCGTGGCTGGCAAGA
GTGGCCTGGAGAAAGAGGTTCAGCGCTTGACCAGCCGAGCTGCCCGTGACTACAAGATCCAGAACCATGGGCATC
GGGTGAGGTGGGGGGGCACAGGTGTCATGTGCACCTTCTTGTCTCAGCAAGAAGAGCTGAGAGAGGGGATCTTGG
AGCCATTGAGGGTGTCATGGAGCTACAGAGGGGAGGGAAAGGTATTTTAAGGTAACAGTGTGGCACAATAGTTAA
GAGCACAGTTTTTGGAGCTAGACCGACATAGGTTCAAATTCTCTTCTGTTGCTTCCTAGTTCTGTAGCCCCAGGT
AAGGGAGTGACTTAACCTCTCTGGACTTCAATTTCCTCATCACTAAAGTAGGGCCAATAATAGCACCCACCTCAT
AGGGAAGATTAAATGACATAATGTATGTGATGCAACTAGCAAAGTACCAGTCCCATAGTAAGTCATGCCCCACAG
TATTTCCACCCACCCCTGTTCTCTGCCTTCCCAACCAGGTACTGCAACGACTGGAGCAGAGGCGGCAGCAGGCTT
CAGAGCGGGAGGCTCCAAGCATAGAACAGAGGTTACAGGAAGTGCGAGAGAGCATCCGCGGGCACAGGTGAGCC
AGGTGAAGGGGGCTGCCCGGCTGGCCCTGCTGCAGGGGGCTGGCTTAGATGTGGAGCGCTGGCTGAAGCCAGCCA
TGACCCAGGCCCAGGATGAGGTGGAGCAGGAGCGGCGGCTCAGTGAGGCTCGGCTGTCCCAGAGGGACCTCTCTC
CAACCGCTGAGGATGCTGAGCTTTCTGACTTTGAGGAATGTGAGGAGACGGGAGAGCTCTTTGAGGAGCCTGCCC
CCCAAGCCCTGGCCACGAGGGGCCCTCCCCTGCCCTGCACACGTGGTATTTCGCTATCAGGCAGGGCGTGAGGATG
AGCTGACAATCACGGAGGGTGAGTGGCTGGAGGTCATAGAGGAGGGAGATGCTGACGAATGGGTCAAGGCTCGGA
ACCAGCACGGCGAGGTAGGCTTTGTCCCTGAGCGATATCTCAACTTCCCGGACCTCTCCCTCCCAGAGAGCAGCC
AAGACAGTGACAATCCCTGCGGGGCAGAGCCCACAGCATTCCTGGCACAGGCCCTGTACAGCTACACCGGACAGA
GTGCAGAGGAGCTGAGCTTCCCTGAGGGGGCACTCATCCGTCTGCTGCCCCCGGGCCCAAGATGGAGTAGATGACG
GCTTCTGGAGGGGAGAATTTGGGGGCCGTGTTGGGGTCTTCCCCTCCCTGCTGGTGGAAGAGCTGCTTGGCCCCC
CAGGGCCACCTGAACTCTCTGACCCTGAACAGATGCTGCCGTCCCCTTCTCCTCCCACCTGCAC
CTACCTCTGTGTTGGATGGGCCCCCTGCACCTGTCCTGCCTGGGGACAAAGCCCTGGACTTCCCTGGGTTCCTGG
ACATGATGGCACCTCGACTCAGGCCGATGCGTCCACCACCTCCCCCGCCGGCTAAAGCCCCGGATCCTGGCCACC
CAGATCCCCTCACCTGAAGGCCAGGGAAGCCTTGACCCCCAGTGATGCTGCTGTCCCTATCTTCAAGCTGTCAGA
CCACACCATCAATGATCCAGAGCAACACAGCCAAAAGCTGGAATCGCCCTTATTTCCACCCTCACCTCCAAGGGT
GGAAACTTGCCCCTTCCCATTTCTAGAGCTGGAACCCACTCCTTTTTTTCCCATTGTTCTATCATCTCTAGGACC
GGAACTACTACCTTCTCTTCTGTCATGACCCTATCTAGGGTGGTGAAATGCCTGAAATCTCTGGGGCTGGAAACC
ATCCATCAAGGTCTCTAGTAGTTCTGGCCCACCTCTTTCCCCACCCTGGCTCCATGACCCACCCCACTCTGGATG
CCAGGGTCACTGGGGTTGGGCTGGGGAGAGGAACAGGCCTTGGGAATCAGGAGCTGGAGCCAGGATGCGAAGCAG
CTGTAATGGTCTGAGCGGATTTATTGACAATGAATAAAAGGGCACGAAGGCCAGGCCAGGGCCTGGGCCTCTTGTG
CTAAGAGGGCAGGGGGCCTACGGTGCTATTGCTTTAGGGGCCCACCACGGGCAGGGGCCTGCTCCCAGCTGCCAC
GCTCTATCATATGGAGCGAGGTGTTGGGGAAGGCGGGGCAGGCAGCCTGTTGCAGGCAGGGGAAGGAGAAGAGAC
TGAGGGGCTGTGACCTCTCCTGAGGCCCCCAGCCTGAGACTGTGCAACTCCAGGTGGAAGTAGAGCTGGTCCCTC
AGCTGGGGGGCAGTGCTGTCTGTCCAGTGGAGGGGAGGGCTTTCACGCCCACCCACCCCCTGGCCCTGCCAGCTGGTAG
TCCATCAGCACAATGAAGGAGACTTGGAGAAGAGGAAGAATAACACTGTTGCTTCCTGTTCAAGCTGTGTCCAGC
TTTTCCCCTGGGGCTCCAGGACCTTCCCTACCTCCACCACCAAACCAAGGGATTTATAGCAAAGGCTAAGCCTGC
AGTTTACTCTGGGGGTTCAGGGAGCCGAAAGGCTTAAATAGTTTAAGTAGGTGATGGGAAGATGAGATTACCTCA
TTTAGGGCTCAGGCAGACTCACCTCACATACTCCCTGCTCCCTGTGGTAGAGACACCTGAGAGAAAGGGGAGGGG
TCAACAATGAGAGACCAGGAGTAGGTCCTATCAGTGCCCCCCAGAGTAGAGAGCAATAAGAGCCCAGCCCAGTGC
AGTCCCGGCTGTGTTTTCCTACCTGGTGATCAGAAGTGTCTGGTTTGCTTGGCTGCCCATTTGCCTCTTGAGTGG
GCAGCCCTGGGCTTGGGCCCCTCCCTCCGGCCCTCAGTGTTGGCTCTGCAGAAGCTCTGGGGTTCCCTTCAAGTG
CACGAGGGGTTAGGCTGCTGTCCCTGAGTCCTCCATTCTGTACTGGGGGGCTGGCTAGGACCTGGGGCTGTGGCC
TCTCAGGGGGCAGCCTCTCCATGGCAGGCATCCCTGCCTTGGGCTGCCCTCCCCCAGACCCCTGACCACCCCCTG
GGTCCTGTCCCCCACCAGAGCCCCAGCTCCTGTCTGTGGGGGAGCCATCACGGTGTTCGTGCAGTCCATAGCGCT
TCTCAATGTGTGTCACCCGGAACCTGGGAGGGGAGGGAACACTGGGGTTTAGGACCACAACTCAGAGGCTGCTTG
GCCCTCCCCTCTGACCAGGGACATCCTGAGTTTGGTGGCTACTTCCCTCTGGCCTAAGGTAGGGGAGGCCTTCTC
AGATTGTGGGGCACATTGTGTAGCCTGACTTCTGCTGGAGCTCCCAGTCCAGGAGGAAAGAGCCAAGGCCCACTT
TTGGGATCAGGTGCCTGATCACTGGGCCCCCTACCTCAGCCCCCCTTTCCCTGGAGCACCTGCCCCACCTGCCCA
CAGAGAACACAGTGGTCTCCCCTGTCCGGGGGCGGCTTTTTCCTTCCTTGGAGCGTCCCTGACGGACAAGTGGAG
GCCTCTTGCTGCGGCTGCAATGGATGCAAGGGGCTGCAGAGCCCAGGTGCACTGTGTGATGATGGGAGGGGGCTC
CGTCCTGCAGGCTGGAGGTGGCATCCACACTGGACAGCAGGAGGAGGGGAGTGAGGGTAACATTTCCATTTCCCT
TCATGTTTTGTTTCTTACGTTCTTTCAGCATGCTCCTTAAAACCCCAGAAGCCCCAATTTCCCCAAGCCCCATTT
TTTCTTGTCTTTATCTAATAAACTCAATATTAAG
```

FIGURE 178

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73401
><subunit 1 of 1, 370 aa, 1 stop
><MW: 40685, pI: 4.53, NX(S/T): 0

MQLAKYQSHSKSCPTVFPPTPVLCLPNQVLQRLEQRRQQASEREAPSIEQRLQEVRESIRRA
QVSQVKGAARLALLQGAGLDVERWLKPAMTQAQDEVEQERRLSEARLSQRDLSPTAEDAELS
DFEECEETGELFEEPAPQALATRALPCPAHVVFRYQAGREDELTITEGEWLEVIEEGDADEW
VKARNQHGEVGFVPERYLNFPDLSLPESSQDSDNPCGAEPTAFLAQALYSYTGQSAEELSFP
EGALIRLLPRAQDGVDDGFWRGEFGGRVGVFPSLLVEELLGPPGPPELSDPEQMLPSPSPPS
FSPPAPTSVLDGPPAPVLPGDKALDFPGFLDMMAPRLRPMRPPPPPPAKAPDPGHPDPLT

FIGURE 179A

```
CACAGGGAGACCCACAGACACATATGCACGAGAGAGACAGAGGAGGAAAGAGACAGAGACAAAGGCACAGCGGAA
GAAGGCAGAGACAGGGCAGGCACAGAAGCGGCCCAGACAGAGTCCTACAGAGGGGAGAGGCCAGAGAAGCTGCAGA
AGACACAGGCAGGGAGAGACAAAGATCCAGGAAAGGAGGGCTCAGGAGGAGAGTTTGGAGAAGCCAGACCCCTGG
GCACCTCTCCCAAGCCCAAGGACTAAGTTTTCTCCATTTCCTTTAACGGTCCTCAGCCCTTCTGAAAACTTTGCC
TCTGACCTTGGCAGGAGTCCAAGCCCCCAGGCTACAGAGAGGAGCTTTCCAAAGCTAGGGTGTGGAGGACTTGGT
GCCCTAGACGGCCTCAGTCCCTCCCAGCTGCAGTACCAGTGCCATGTCCCAGACAGGCTCGCATCCCGGGAGGGG
CTTGGCAGGGCGCTGGCTGTGGGGAGCCCAACCCTGCCTCCTGCTCCCCATTGTGCCGCTCTCCTGGCTGGTGTG
GCTGCTTCTGCTACTGCTGGCCTCTCTCCTGCCCTCAGCCCGGCTGGCCAGCCCCCTCCCCGGGAGGAGGAGAT
CGTGTTTCCAGAGAAGCTCAACGGCAGCGTCCTGCCTGGCTCGGGCGCCCCTGCCAGGCTGTTGTGCCGCTTGCA
GGCCTTTGGGGAGACGCTGCTACTAGAGCTGGAGCAGGACTCCGGTGTGCAGGTCGAGGGGCTGACAGTGCAGTA
CCTGGGCCAGGCGCCTGAGCTGCTGGGTGGAGCAGAGCCTGGCACCTACCTGACTGGCACCATCAATGGAGATCC
GGAGTCGGTGGCATCTCTGCACTGGGATGGGGGAGCCCTGTTAGGCGTGTTACAATATCGGGGGGCTGAACTCCA
CCTCCAGCCCCTGGAGGGAGGCACCCCTAACTCTGCTGGGGGACCTGGGGCTCACATCCTACGCCGGAAGAGTCC
TGCCAGCGGTCAAGGTCCCATGTGCAACGTCAAGGCTCCTCTTGGAAGCCCCAGCCCCAGACCCCGAAGAGCCAA
GCGCTTTGCTTCACTGAGTAGATTTGTGGAGACACTGGTGGTGGCAGATGACAAGATGGCCGCATTCCACGGTGC
GGGGCTAAAGCGCTACCTGCTAACAGTGATGGCAGCAGCAGCCAAGGCCTTCAAGCACCCAAGCATCCGCAATCC
TGTCAGCTTGGTGGTGACTCGGCTAGTGATCCTGGGGTCAGGCGAGGAGGGGCCCCAAGTGGGGCCCAGTGCTGC
CCAGACCCTGCGCAGCTTCTGTGCCTGGCAGCGGGGCCTCAACACCCCTGAGGACTCGGGCCCTGACCACTTTGA
CACAGCCATTCTGTTTACCCGTCAGGACCTGTGTGGAGTCTCCACTTGCGACACGCTGGGTATGGCTGATGTGGG
CACCGTCTGTGACCCGGCTCGGAGCTGTGCCATTGTGGAGGATGATGGGCTCCAGTCAGCCTTCACTGCTGCTCA
TGAACTGGGTCATGTCTTCAACATGCTCCATGACAACTCCAAGCCATGCATCAGTTTGAATGGGCCTTTGAGCAC
CTCTCGCCATGTCATGGCCCCTGTGATGGCTCATGTGGATCCTGAGGAGCCCTGGTCCCCCTGCAGTGCCCGCTT
CATCACTGACTTCCTGGACAATGGCTATGGGCACTGTCTCTTAGACAAACCAGAGGCTCCATTGCATCTGCCTGT
GACTTTCCCTGGCAAGGACTATGATGCTGACCGCCAGTGCCAGCTGACCTTCGGGCCCGACTCACGCCATTGTCC
ACAGCTGCCGCCGCCCTGTGCTGCCCTCTGGTGCTCTGGCCACCTCAATGGCCATGCCATGTGCCAGACCAAACA
CTCGCCCTGGGCCGATGGCACACCCTGCGGGCCCGCACAGGCCTGCATGGGTGGTCGCTGCCTGCCTCCACATGGACCA
GCTCCAGGACTTCAATATTCCACAGGCTGGTGGCTGGGGTCCTTGGGGACCATGGGGTGACTGCTCTCGGACCTG
TGGGGGTGGTGTCCAGTTCTCCTCCCGAGACTGCACGAGGCCTGTCCCCCGGAATGGTGGCAAGTACTGTGAGGG
CCGCCGTACCCGCTTCCGCTCCTGCAACACTGAGGACTGCCCAACTGGCTCAGCCCTGACCTTCCGCGAGGAGCA
GTGTGCTGCCTACAACCACCGCACCGACCTCTTCAAGAGCTTCCCAGGGCCCATGGACTGGGTTCCTCGCTACAC
AGGCGTGGCCCCCCAGGACCAGTGCAAACTCACCTGCCAGGCCCGGGCACTGGGCTACTACTATGTGCTGGAGCC
ACGGGTGGTAGATGGGACCCCCTGTTCCCCGGACAGCTCCTCGGTCTGTGTCCAGGGCCGATGCATCCATGCTGG
CTGTGATCGCATCATTGGCTCCAAGAAGAAGTTTGACAAGTGCATGGTGTGCGGAGGGGACGGTTCTGGTTGCAG
CAAGCAGTCAGGCTCCTTCAGGAAATTCAGGTACGGATACAACAATGTGGTCACTATCCCCGCGGGGGCCACCCA
CATTCTTGTCCGGCAGCAGGGAAACCCTGGCCACCGGAGCATCTACTTGGCCCTGAAGCTGCCAGATGGCTCCTA
TGCCCTCAATGGTGAATACACGCTGATGCCCTCCCCCACAGATGTGGTACTGCCTGGGGCAGTCAGCTTGCGCTA
CAGCGGGGCCACTGCAGCCTCAGAGACACTGTCAGGCCATGGGCCACTGGCCCAGCCTTTGACACTGCAAGTCCT
AGTGGCTGGCAACCCCCAGGACACACGCCTCCGATACAGCTTCTTCGTGCCCCGGCCGACCCCTTCAACGCCACG
CCCCACTCCCCAGGACTGGCTGCACCGAAGAGCACAGATTCTGGAGATCCTTCGGCGGCGCCCCTGGGCGGGCAG
GAAATAACCTCACTATCCCGGCTGCCCTTTCTGGGCACCGGGGCCTCGGACTTAGCTGGGAGAAAGAGAGAGCTT
CTGTTGCTGCCTCATGCTAAGACTCAGTGGGGAGGGGCTGTGGGCGTGAGACCTGCCCCTCCTCTCTGCCCTAAT
GCGCAGGCTGGCCCTGCCCTGGTTTCCTGCCCTGGGAGGCAGTGATGGGTTAGTGGATGGAAGGGGCTGACAGAC
AGCCCTCCATCTAAACTGCCCCCTCTGCCCTGCGGGTCACAGGAGGGGAGGGGGAAGGCAGGGAGGGCCTGGGCCC
CAGTTGTATTTATTTAGTATTTATTCACTTTTATTTAGCACCAGGGAAGGGGACAAGGACTAGGGTCCTGGGGAA
CCTGACCCCTGACCCCTCATAGCCCTCACCCTGGGGCTAGGAAATCCAGGGTGGTGGTGATAGGTATAAGTGGTG
TGTGTATGCGTGTGTGTGTGTGTGTGTGAAAATGTGTGTGTGCTTATGTATGAGGTACAACCTGTTCTGCTTTCCTC
TTCCTGAATTTTATTTTTTGGGAAAAGAAAAGTCAAGGGTAGGGTGGGCCTTCAGGGAGTGAGGGATTATCTTTT
TTTTTTTTTCTTTCTTTCTTTCTTTTTTTTTTTTTGAGACAGAATCTCGCTCTGTCGCCCAGGCTGGAGTGCAATG
GCACAATCTCGGCTCACTGCATCCTCCGCCTCCCGGGTTCAAGTGATTCTCATGCCTCAGCCTCCTGAGTAGCTG
GGATTACAGGCTCCTGCCACCACGCCCAGCTAATTTTTGTTTTGTTTTGTTTGGAGACAGAGTCTCGCTATTGTC
ACCAGGGCTGGAATGATTTCAGCTCACTGCAACCTTCGCCACCTGGGTTCCAGCAATTCTCCTGCCTCAGCCTCC
CGAGTAGCTGAGATTATAGGCACCTACCACCACGCCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCAC
CATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTTAGGTGATCCACTCGCCTTCATCTCCCAAAGTGCTGGGATT
ACAGGCGTGAGCCACCGTGCCTGGCCACGCCCAACTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGT
TGGCCAGGCTGCTCTTGAACTCCTGACCTCAGGTAATCGACCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG
TGTGAGCCACCACGCCCGGTACATATTTTTTAAATTGAATTCTACTATTTATGTGATCCTTTTGGAGTCAGACAG
```

FIGURE 179B

ATGTGGTTGCATCCTAACTCCATGTCTCTGAGCATTAGATTTCTCATTTGCCAATAATAATACCTCCCTTAGAAG
TTTGTTGTGAGGATTAAATAATGTAAATAAAGAACTAGCATAACACTCAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAAA

FIGURE 180

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73492
><subunit 1 of 1, 837 aa, 1 stop
><MW: 90167, pI: 8.39, NX(S/T): 1
MSQTGSHPGRGLAGRWLWGAQPCLLLPIVPLSWLVWLLLLLLASLLPSARLASPLPREEEIV
FPEKLNGSVLPGSGAPARLLCRLQAFGETLLLELEQDSGVQVEGLTVQYLGQAPELLGGAEP
GTYLTGTINGDPESVASLHWDGGALLGVLQYRGAELHLQPLEGGTPNSAGGPGAHILRRKSP
ASGQGPMCNVKAPLGSPSPRPRRAKRFASLSRFVETLVVADDKMAAFHGAGLKRYLLTVMAA
AAKAFKHPSIRNPVSLVVTRLVILGSGEEGPQVGPSAAQTLRSFCAWQRGLNTPEDSGPDHF
DTAILFTRQDLCGVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAPTAAHELGHVFNMLHD
NSKPCISLNGPLSTSRHVMAPVMAHVDPEEPWSPCSARFITDFLDNGYGHCLLDKPEAPLHL
PVTFPGKDYDADRQCQLTFGPDSRHCPQLPPPCAALWCSGHLNGHAMCQTKHSPWADGTPCG
PAQACMGGRCLHMDQLQDFNIPQAGGWGPWGPWGDCSRTCGGGVQFSSRDCTRPVPRNGGKY
CEGRRTRFRSCNTEDCPTGSALTFREEQCAAYNHRTDLFKSFPGPMDWVPRYTGVAPQDQCK
LTCQARALGYYYVLEPRVVDGTPCSPDSSSVCVQGRCIHAGCDRIIGSKKKFDKCMVCGGDG
SGCSKQSGSFRKFRYGYNNVVTIPAGATHILVRQQGNPGHRSIYLALKLPDGSYALNGEYTL
MPSPTDVVLPGAVSLRYSGATAASETLSGHGPLAQPLTLQVLVAGNPQDTRLRYSFFVPRPT
PSTPRPTPQDWLHRRAQILEILRRRPWAGRK
```

Important features of the protein:

Signal peptide:

amino acids 1-48

N-glycosylation site.

amino acids 68-71

Glycosaminoglycan attachment site

amino acids 188-191, 772-775

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 182-185

Tyrosine kinase phosphorylation site.

amino acids 730-736

N-myristoylation sites.

amino acids 5-10, 19-24, 121-126, 125-130, 130-135, 147-152, 167-

172, 168-173, 174-179, 323-328, 352-357, 539-544, 555-560, 577-

582, 679-684, 682-687, 763-768

Amidation sites.

amino acids 560-563, 834-837

Leucine zipper pattern.

amino acids 17-38, 24-45

Neutral zinc metallopeptidases, zinc-binding region signature.
amino acids 358-367

183/270

FIGURE 181

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACC**ATG**G

CAAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCC

AAGAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTTTGGTATCCTGGCCCTAACTCT

AATTGTCCTGTTTTGGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAAGCCTATGACA

TGGAGCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTG

ACCAGAACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTT

TAAAAACGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAAACTCAGA

TTAAAGTGATTCCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACC

ACAACTTTCTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGA

TTTTCTTAAAAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATC

CCACTCTAATATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTT

CCTGCCAACGAAAAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGT

AGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTG

AAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGT

CGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATA

CTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTGTAACTGGTGGGTGGCCC

GCATGCTGGGGAGGGTC**TAA**TAGGAGGTTTGAGCTCAAATGCTTAAACTGCTGGCAACATAT

AATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCCTGGTAGCCAGCT

CTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACC

AAAAAAAAAAAAAAAAAAA

FIGURE 182

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73727

><subunit 1 of 1, 317 aa, 1 stop

><MW: 37130, pI: 5.18, NX(S/T): 3

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPKKAY

DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKT

QIKVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKNSKILEICDNVTMYWI

NPTLISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEEELPINDY

TENGIEFDPMLDERGYCCIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWV

ARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 183

GCGGAACTGGCTCCGGCTGGCACCTGAGGAGCGGCGTGACCCCGAGGGCCCAGGGAGCTGCC

CGGCTGGCCTAGGCAGGCAGCCGCACCATGGCCAGCACGGCCGTGCAGCTTCTGGGCTTCCT

GCTCAGCTTCCTGGGCATGGTGGGCACGTTGATCACCACCATCCTGCCGCACTGGCGGAGGA

CAGCGCACGTGGGCACCAACATCCTCACGGCCGTGTCCTACCTGAAAGGGCTCTGGATGGAG

TGTGTGTGGCACAGCACAGGCATCTACCAGTGCCAGATCTACCGATCCCTGCTGGCGCTGCC

CCAAGACCTCCAGGCTGCCCGCGCCCTCATGGTCATCTCCTGCCTGCTCTCGGGCATAGCCT

GCGCCTGCGCCGTCATCGGGATGAAGTGCACGCGCTGCGCCAAGGGCACACCCGCCAAGACC

ACCTTTGCCATCCTCGGCGGCACCCTCTTCATCCTGGCCGGCCTCCTGTGCATGGTGGCCGT

CTCCTGGACCACCAACGACGTGGTGCAGAACTTCTACAACCCGCTGCTGCCCAGCGGCATGA

AGTTTGAGATTGGCCAGGCCCTGTACCTGGGCTTCATCTCCTCGTCCCTCTCGCTCATTGGT

GGCACCCTGCTTTGCCTGTCCTGCCAGGACGAGGCACCCTACAGGCCCTACCAGGCCCCGCC

CAGGGCCACCACGACCACTGCAAACACCGCACCTGCCTACCAGCCACCAGCTGCCTACAAAG

ACAATCGGGCCCCCTCAGTGACCTCGGCCACGCACAGCGGGTACAGGCTGAACGACTACGTG

TGAGTCCCCACAGCCTGCTTCTCCCCTGGGCTGCTGTGGGCTGGGTCCCCGGCGGGACTGTC

AATGGAGGCAGGGGTTCCAGCACAAAGTTTACTTCTGGGCAATTTTTGTATCCAAGGAAATA

ATGTGAATGCGAGGAAATGTCTTTAGAGCACAGGGACAGAGGGGGAAATAAGAGGAGGAGAA

AGCTCTCTATACCAAAGACTGAAAAAAAAAATCCTGTCTGTTTTTGTATTTATTATATATAT

TTATGTGGGTGATTTGATAACAAGTTTAATATAAAGTGACTTGGGAGTTTGGTCAGTGGGGT

TGGTTTGTGATCCAGGAATAAACCTTGCGGATGTGGCTGTTTATGAAAAAAAAAAAAA

186/270

FIGURE 184

MASTAVQLLGFLLSFLGMVGTLITTILPHWRRTAHVGTNILTAVSYLKGLWMECVWHSTGIY

QCQIYRSLLALPQDLQAARALMVISCLLSGIACACAVIGMKCTRCAKGTPAKTTFAILGGTL

FILAGLLCMVAVSWTTNDVVQNFYNPLLFSGMKFEIGQALYLGFISSSLSLIGGTLLCLSCQ

DEAPYRPYQAPPRATTTTANTAPAYQPPAAYKDNRAPSVTSATHSGYRLNDYV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-103, 115-141, 160-182

FIGURE 185

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCG
CAGGCGGCAGGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCCT
CCTGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCC
AGGACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGC
GTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCCTATTTCACCATCCTGGGACTTCCAGC
CATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCC
TGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCC
AACATGACACTGACCTCCGGGATCATGTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGT
GTCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCG
GCATGGGTGGGATGGTGCAGACTGTTCAGACCAGGTACACATTTGGTGCGGCTCTGTTCGTG
GGCTGGGTCGCTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGG
CCTGGCACCAGAAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTG
CCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAG
ATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTA
TGTGTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAAACTCCCGGAGAGCTCACCCAAAA
AACAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGC
CTCGATTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCC
ACCATAAAACAGCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTAT
TTCTTTTTTTTAAATATAACTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCAC
ATTTTGATGATTTAGACAGACTCCCCCTCTTCCTCCTAGTCAATAAACCCATTGATGATCTA
TTTCCCAGCTTATCCCCAAGAAAACTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTTT
CTGCTGTTTGAATTTTGTCTCCCCACCCCCAACTTGGCTAGTAATAAACACTTACTGAAGAA
GAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTG
TGATCTTAAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTG
CTGTTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAG
TCCTCTTTCTGTCGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAT
TTAAGTCCTAAATATAGTTAAAATAAATAATGTTTTAGTAAAATGATACACTATCTCTGTGA
AATAGCCTCACCCCTACATGTGGATAGAAGGAAATGAAAAAATAATTGCTTTGACATTGTCT
ATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCCATGAAAAGCTCACACCTGTAATC
CTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGTTCGAGACTAGCCTG
GGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCATGGTG
GCATACACCTGTAGTCCCAGCATTCCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGG
AGGTTGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGA
TCCTGTCTAAAAAAATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAA
ACTAATTCTTTAA

FIGURE 186

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73734
><subunit 1 of 1, 261 aa, 1 stop
><MW: 27856, pI: 8.50, NX(S/T): 1
MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTE
CRPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIM
FIVSGLCAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTL
IGGVMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTE
DEVQSYPSKHDYV
```

Signal peptide:
amino acids 1-23

Transmembrane domains:
amino acids 81-100, 121-141, 173-194

FIGURE 187

GGAAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCG
GAGTCCAGCTGGCTAAAACTCATCCCAGAGGATA**ATG**GCAACCCATGCCTTAGAAATCGCTG
GGCTGTTTCTTGGTGGTGTTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGG
AGAGTGTCGGCCTTCATTGAAAACAACATCGTGGTTTTTGAAAACTTCTGGGAAGGACTGTG
GATGAATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGG
CTCTTTCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTC
TTGGCTTTCATGATGGCCATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAA
GGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGC
TCATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAAT
GTTGCCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGCT
GATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAGAGCAGTAGCTACA
GATACTCGATACCTTCCCATCGCACAACCCAAAAAAGTTATCACACCGGAAAGAAGTCACCG
AGCGTCTACTCCAGAAGTCAGTATGTG**TAG**TTGTGTATGTTTTTTTAACTTTACTATAAAGC
CATGCAAATGACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAAGAAACTTTGATTTA
CTGTTCTTAACTGCCTAATCTTAATTACAGGAACTGTGCATCAGCTATTTATGATTCTATAA
GCTATTTCAGCAGAATGAGATATTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAAT
TTGTTTTCTAAGGTGGTTCAAGCATCTACTCTTTTTATCATTTACTTCAAAATGACATTGCT
AAAGACTGCATTATTTTACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAG
TGTAACATTTATATCTCACATAGAGACATGCTTATATGGTTTTATTTAAAATGAAATGCCAG
TCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAAATCATGGATAGGGTTG
AAGAAGGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCATTTATAATGA
AGATTAAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTG
TTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTT
TTCTTGTGTATTAAATTAACATTTTTAAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCA
AACTGCTTTTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATG
GTTTTAGGAAAGTGAAAATATTTTTGTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAA
GAAATCATATATGTATGGATATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATC
AATATAAATAAAAGAGCAGAAAAATATGTCTTGGTTTTCATTTGCTTACCAAAAAAACAACA
ACAAAAAAAGTTGTCCTTTGAGAACTTCACCTGCTCCTATGTGGGTACCTGAGTCAAAATTG
TCATTTTTGTTCTGTGAAAAATAAATTTCCTTCTTGTACCATTTCTGTTTAGTTTTACTAAA
ATCTGTAAATACTGTATTTTTCTGTTTATTCCAAATTTGATGAAACTGACAATCCAATTTGA
AAGTTTGTGTCGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTTTATACATTTATA
TTAATAAATTGTACATTTTTCTAATT

190/270

FIGURE 188

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73735

><subunit 1 of 1, 225 aa, 1 stop

><MW: 24845, pI: 9.07, NX(S/T): 0

MATHALEIAGLFLGGVGMVGTVAVTVMPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRM

QCKIYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGI

IFIITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVF

CCNEKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 189

TCGCC**ATG**GCCTCTGCCGGAATGCAGATCCTGGGAGTCGTCCTGACACTGCTGGGCTGGGTG

AATGGCCTGGTCTCCTGTGCCCTGCCCATGTGGAAGGTGACCGCTTTCATCGGCAACAGCAT

CGTGGTGGCCCAGGTGGTGTGGGAGGGCCTGTGGATGTCCTGCGTGGTGCAGAGCACCGGCC

AGATGCAGTGCAAGGTGTACGACTCACTGCTGGCGCTGCCACAGGACCTGCAGGCTGCACGT

GCCCTCTGTGTCATCGCCCTCCTTGTGGCCCTGTTCGGCTTGCTGGTCTACCTTGCTGGGGC

CAAGTGTACCACCTGTGTGGAGGAGAAGGATTCCAAGGCCCGCCTGGTGCTCACCTCTGGGA

TTGTCTTTGTCATCTCAGGGGTCCTGACGCTAATCCCCGTGTGCTGGACGGCGCATGCCATC

ATCCGGGACTTCTATAACCCCCTGGTGGCTGAGGCCCAAAAGCGGGAGCTGGGGGCCTCCCT

CTACTTGGGCTGGGCGGCCTCAGGCCTTTTGTTGCTGGGTGGGGGGTTGCTGTGCTGCACTT

GCCCCTCGGGGGGGTCCCAGGGCCCCAGCCATTACATGGCCCGCTACTCAACATCTGCCCCT

GCCATCTCTCGGGGGCCCTCTGAGTACCCTACCAAGAATTACGTC**TGA**CGTGGAGGGGAATG

GGGGCTCCGCTGGCGCTAGAGCCATCCAGAAGTGGCAGTGCCCAACAGCTTTGGGATGGGTT

CGTACCTTTTGTTTCTGCCTCCTGCTATTTTTCTTTTGACTGAGGATATTTAAAATTCATTT

GAAAACTGAGCCAAGGTGTTGACTCAGACTCTCACTTAGGCTCTGCTGTTTCTCACCCTTGG

ATGATGGAGCCAAAGAGGGGATGCTTTGAGATTCTGGATCTTGACATGCCCATCTTAGAAGC

CAGTCAAGCTATGGAACTAATGCGGAGGCTGCTTGCTGTGCTGGCTTTGCAACAAGACAGAC

TGTCCCCAAGAGTTCCTGCTGCTGCTGGGGGCTGGGCTTCCCTAGATGTCACTGGACAGCTG

CCCCCCATCCTACTCAGGTCTCTGGAGCTCCTCTCTTCACCCCTGGAAAAACAAATCATCTG

TTAACAAAGGACTGCCCACCTCCGGAACTTCTGACCTCTGTTTCCTCCGTCCTGATAAGACG

TCCACCCCCCAGGGCCAGGTCCCAGCTATGTAGACCCCCGCCCCCACCTCCAACACTGCACC

CTTCTGCCCTGCCCCCCTCGTCTCACCCCCTTTACACTCACATTTTTATCAAATAAAGCATG

TTTTGTTAGTGCA

FIGURE 190

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73736

><subunit 1 of 1, 220 aa, 1 stop

><MW: 23292, pI: 8.43, NX(S/T): 0

MASAGMQILGVVLTLLGWVNGLVSCALPMWKVTAFIGNSIVVAQVVWEGLWMSCVVQSTGQM

QCKVYDSLLALPQDLQAARALCVIALLVALFGLLVYLAGAKCTTCVEEKDSKARLVLTSGIV

FVISGVLTLIPVCWTAHAIIRDFYNPLVAEAQKRELGASLYLGWAASGLLLLGGGLLCCTCP

SGGSQGPSHYMARYSTSAPAISRGPSEYPTKNYV

Transmembrane domains:

amino acids 8-30 (type II), 82-102, 121-140, 166-186

FIGURE 191

GCCAAGGAGAACATCATCAAAGACTTCTCTAGACTCAAAAGGCTTCCACGTTCTACATCTTG

AGCATCTTCTACCACTCCGAATTGAACCAGTCTTCAAAGTAAAGGCAATGGCATTTTATCCC

TTGCAAATTGCTGGGCTGGTTCTTGGGTTCCTTGGCATGGTGGGGACTCTTGCCACAACCCT

TCTGCCTCAGTGGTGGAGTATCAGCTTTTGTTGGCAGCAACATTATTGTCTTTGAGAGGCTC

TGGGAAGGGCTCTGG**ATG**AATTGCATCCGACAAGCCAGGGTCCGGTTGCAATGCAAGTTCTA

TAGCTCCTTGTTGGCTCTCCCGCCTGCCCTGGAAACAGCCCGGGCCCTCATGTGTGTGGCTG

TTGCTCTCTCCTTGATCGCCCTGCTTATTGGCATCTGTGGCATGAAGCAGGTCCAGTGCACA

GGCTCTAACGAGAGGGCCAAAGCATACCTTCTGGGAACTTCAGGAGTCCTCTTCATCCTGAC

GGGTATCTTCGTTCTGATTCCGGTGAGCTGGACAGCCAATATAATCATCAGAGATTTCTACA

ACCCAGCCATCCACATAGGTCAGAAACGAGAGCTGGGAGCAGCACTTTTCCTTGGCTGGGCA

AGCGCTGCTGTCCTCTTCATTGGAGGGGGTCTGCTTTGTGGATTTTGCTGCTGCAACAGAAA

GAAGCAAGGGTACAGATATCCAGTGCCTGGCTACCGTGTGCCACACACAGATAAGCGAAGAA

ATACGACAATGCTTAGTAAGACCTCCACCAGTTATGTC**TAA**TGCCTCCTTTTGGCTCCAAGT

ATGGACTATGGTCAATGTTTTTTATAAAGTCCTGCTAGAAACTGTAAGTATGTGAGGCAGGA

GAACTTGCTTTATGTCTAGATTTACATTGATACGAAAGTTTCAATTTGTTACTGGTGGTAGG

AATGAAAATGACTTACTTGGACATTCTGACTTCAGGTGTATTAAATGCATTGACTATTGTTG

GACCCAATCGCTGCTCCAATTTTCATATTCTAAATTCAAGTATACCCATAATCATTAGCAAG

TGTACAATGATGGACTACTTATTACTTTTTGACCATCATGTATTATCTGATAAGAATCTAAA

GTTGAAATTGATATTCTATAACAATAAAACATATACCTATTCTA

FIGURE 192

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73737

><subunit 1 of 1, 173 aa, 1 stop

><MW: 18938, pI: 9.99, NX(S/T): 1

MNCIRQARVRLQCKFYSSLLALPPALETARALMCVAVALSLIALLIGICGMKQVQCTGSNER

AKAYLLGTSGVLFILTGIFVLIPVSWTANIIIRDFYNPAIHIGQKRELGAALFLGWASAAVL

FIGGGLLCGFCCCNRKKQGYRYPVPGYRVPHTDKRRNTTMLSKTSTSYV

Important features of the protein:

Transmembrane domains:

amino acids 31-51, 71-90, 112-133

N-glycosylation site.

amino acids 161-164

FIGURE 193

AGTGACAATCTCAGAGCAGCTTCTACACCACAGCCATTTCCAGC**ATG**AAGATCACTGGGGGT

CTCCTTCTGCTCTGTACAGTGGTCTATTTCTGTAGCAGCTCAGAAGCTGCTAGTCTGTCTCC

AAAAAAAGTGGACTGCAGCATTTACAAGAAGTATCCAGTGGTGGCCATCCCCTGCCCCATCA

CATACCTACCAGTTTGTGGTTCTGACTACATCACCTATGGGAATGAATGTCACTTGTGTACC

GAGAGCTTGAAAAGTAATGGAAGAGTTCAGTTTCTTCACGATGGAAGTTGC**TAA**ATTCTCCA

TGGACATAGAGAGAAAGGAATGATATTCTCATCATCATCTTCATCATCCCAGGCTCTGACTG

AGTTTCTTTCAGTTTTACTGATGTTCTGGGTGGGGGACAGAGCCAGATTCAGAGTAATCTTG

ACTGAATGGAGAAAGTTTCTGTGCTACCCCTACAAACCCATGCCTCACTGACAGACCAGCAT

TTTTTTTTTTAACACGTCAATAAAAAAATAATCTCCCAGA

196/270

FIGURE 194

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73739

><subunit 1 of 1, 85 aa, 1 stop

><MW: 9232, pI: 7.94, NX(S/T): 0

MKITGGLLLLCTVVYFCSSSEAASLSPKKVDCSIYKKYPVVAIPCPITYLPVCGSDYITYGN

ECHLCTESLKSNGRVQFLHDGSC

Signal peptide:

amino acids 1-19

FIGURE 195

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCCTCGCCCTCCTGCTCGCGCCC

CGCCGCC**ATG**GCTGCCTCCCCCGCGCGGCCTGCTGTCCTGGCCCTGACCGGGCTGGCGCTGC

TCCTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAAATAAACTCAAGCTGATGCTTCAA

AAACGAGAAGCACCTGTTCCAACTAAGACTAAAGTGGCCGTTGATGAGAATAAAGCCAAAGA

ATTCCTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGGACCGGACTCGGCCCGAGGTGC

AGCAGTGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTTGAAGATGACATC

ACCTATTGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGATTACTACCAACGTCA

CTATGATGAAGACTCTGCAATTGGTCCCCGGAGCCCCTACGGCTTTAGGCATGGAGCCAGCG

TCAACTACGATGACTAC**TAA**CCATGACTTGCCACACGCTGTACAAGAAGCAAATAGCGATTC

TCTTCATGTATCTCCTAATGCCTTACACTACTTGGTTTCTGATTTGCTCTATTTCAGCAGAT

CTTTTCTACCTACTTTGTGTGATCAAAAAAGAAGAGTTAAAACAACACATGTAAATGCCTTT

TGATATTTCATGGGAATGCCTCTCATTTAAAAATAGAAATAAAGCATTTTGTTAAAAAGA

FIGURE 196

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73742

><subunit 1 of 1, 148 aa, 1 stop

><MW: 17183, pI: 8.77, NX(S/T): 0

MAASPARPAVLALTGLALLLLLCWGPGGISGNKLKLMLQKREAPVPTKTKVAVDENKAKEFL

GSLKRQKRQLWDRTRPEVQQWYQQFLYMGFDEAKFEDDITYWLNRDRNGHEYYGDYYQRHYD

EDSAIGPRSPYGFRHGASVNYDDY

Signal peptide:

amino acids 1-30

FIGURE 197

CGGCTCGAGCCCGCCCGGAAGTGCCCGAGGGGCCGCGATGGAGCTGGGGGAGCCGGGCGCTC

GGTAGCGCGGCGGGCAAGGCAGGCGCC**ATG**ACCCTGATTGAAGGGGTGGGTGATGAGGTGAC

CGTCCTTTTCTCGGTGCTTGCCTGCCTTCTGGTGCTGGCCCTTGCCTGGGTCTCAACGCACA

CCGCTGAGGGCGGGGACCCACTGCCCCAGCCGTCAGGGACCCCAACGCCATCCCAGCCCAGC

GCAGCCATGGCAGCTACCGACAGCATGAGAGGGGAGGCCCCAGGGGCAGAGACCCCCAGCCT

GAGACACAGAGGTCAAGCTGCACAGCCAGAGCCCAGCACGGGGTTCACAGCAACACCGCCAG

CCCCGGACTCCCCGCAGGAGCCCCTCGTGCTACGGCTGAAATTCCTCAATGATTCAGAGCAG

GTGGCCAGGGCCTGGCCCCACGACACCATTGGCTCCTTGAAAAGGACCCAGTTTCCCGGCCG

GGAACAGCAGGTGCGACTCATCTACCAAGGGCAGCTGCTAGGCGACGACACCCAGACCCTGG

GCAGCCTTCACCTCCCTCCCAACTGCGTTCTCCACTGCCACGTGTCCACGAGAGTCGGTCCC

CCAAATCCCCCCTGCCCGCCGGGGTCCGAGCCCGGCCCCTCCGGGCTGGAAATCGGCAGCCT

GCTGCTGCCCCTGCTGCTCCTGCTGTTGCTGCTGCTCTGGTACTGCCAGATCCAGTACCGGC

CCTTCTTTCCCCTGACCGCCACTCTGGGCCTGGCCGGCTTCACCCTGCTCCTCAGTCTCCTG

GCCTTTGCCATGTACCGCCCG**TAG**TGCCTCCGCGGGCGCTTGGCAGCGTCGCCGGCCCCTCC

GGACCTTGCTCCCCGCGCCGCGGCGGGAGCTGCTGCCTGCCCAGGCCCGCCTCTCCGGCCTG

CCTCTTCCCGCTGCCCTGGAGCCCAGCCCTGCGCCGCAGAGGACTCCCGGGACTGGCGGAGG

CCCCGCCCTGCGACCGCCGGGGCTCGGGGCCACCTCCCGGGGCTGCTGAACCTCAGCCCGCA

CTGGGAGTGGGCTCCTCGGGGTCGGGCATCTGCTGTCGCTGCCTCGGCCCCGGGCAGAGCCG

GGCCGCCCCGGGGGCCCGTCTTAGTGTTCTGCCGGAGGACCCAGCCGCCTCCAATCCCTGAC

AGCTCCTTGGGCTGAGTTGGGGACGCCAGGTCGGTGGGAGGCTGGTGAAGGGGAGCGGGGAG

GGGCAGAGGAGTTCCCCGGAACCCGTGCAGATTAAAGTAACTGTGAAGTTTTAAAAAAAAAA

AAAAAAAA

FIGURE 198

MTLIEGVGDEVTVLFSVLACLLVLALAWVSTHTAEGGDPLPQPSGTPTPSQPSAAMAATDSM

RGEAPGAETPSLRHRGQAAQPEPSTGFTATPPAPDSPQEPLVLRLKFLNDSEQVARAWPHDT

IGSLKRTQFPGREQQVRLIYQGQLLGDDTQTLGSLHLPPNCVLHCHVSTRVGPPNPPCPPGS

EPGPSGLEIGSLLLPLLLLLLLLLWYCQIQYRPFFPLTATLGLAGFTLLLSLLAFAMYRP

Signal peptide:
amino acids 1-31

Transmembrane domain:
amino acids 195-217

FIGURE 199

GAGATTGGAAACAGCCAGGTTGGAGCAGTGAGTGAGTAAGGAAACCTGGCTGCCCTCTCCAG

ATTCCCCAGGCTCTCAGAGAAGATCAGCAGAAAGTCTGCAAGACCCTAAGAACCATCAGCCC

TCAGCTGCACCTCCTCCCCTCCAAGG<u>ATG</u>ACAAAGGCGCTACTCATCTATTTGGTCAGCAGC

TTTCTTGCCCTAAATCAGGCCAGCCTCATCAGTCGCTGTGACTTGGCCCAGGTGCTGCAGCT

GGAGGACTTGGATGGGTTTGAGGGTTACTCCCTGAGTGACTGGCTGTGCCTGGCTTTTGTGG

AAAGCAAGTTCAACATATCAAAGATAAATGAAAATGCGGATGGAAGCTTTGACTATGGCCTC

TTCCAGATCAACAGCCACTACTGGTGCAACGATTATAAGAGTTACTCGGAAAACCTTTGCCA

CGTAGACTGTCAAGATCTGCTGAATCCCAACCTTCTTGCAGGCATCCACTGCGCAAAAAGGA

TTGTGTCCGGAGCACGGGGGATGAACAACTGGGTAGAATGGAGGTTGCACTGTTCAGGCCGG

CCACTCTCCTACTGGCTGACAGGATGCCGCCTGAGA<u>**TGA**</u>AACAGGGTGCGGGTGCACCGTGG

AGTCATTCCAAGACTCCTGTCCTCACTCAGGGATTCTTCATTTCTTCTTCCTACTGCCTCCA

CTTCATGTTATTTTCTTCCCTTCCCATTTACAACTAAAACTGACCAGAGCCCCAGGAATAAA

TGGTTTTCTTGGCTTCCTCCTTACTCCCATCTGGACCCAGTCCCCTGGTTCCTGTCTGTTAT

TTGTAAACTGAGGACCACAATAAAGAAATCTTTATATTTATCG

FIGURE 200

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73746

><subunit 1 of 1, 148 aa, 1 stop

><MW: 16896, pI: 6.05, NX(S/T): 1

MTKALLIYLVSSFLALNQASLISRCDLAQVLQLEDLDGFEGYSLSDWLCLAFVESKFNISKI

NENADGSFDYGLFQINSHYWCNDYKSYSENLCHVDCQDLLNPNLLAGIHCAKRIVSGARGMN

NWVEWRLHCSGRPLSYWLTGCRLR

Signal peptide:
amino acids 1-18

FIGURE 201

```
TCTGACCTGACTGGAAGCGTCCAAAGAGGGACGGCTGTCAGCCCTGCTTGACTGAGAACCCA
CCAGCTCATCCCAGACACCTCATAGCAACCTATTTATACAAAGGGGGAAAGAAACACCTGAG
CAGAATGGAATCATTATTTTTTTCCCAAGGAGAAAACCGGGGTAAAGGGAGGGAAGCAATTC
AATTTGAAGTCCCTGTGAATGGGCTTTCAGAAGGCAATTAAAGAAATCCACTCAGAGAGGAC
TTGGGGTGAAACTTGGGTCCTGTGGTTTTCTGATTGTAAGTGGAAGCAGGTCTTGCACACGC
TGTTGGCAAATGTCAGGACCAGGTTAAGTGACTGGCAGAAAAACTTCCAGGTGGAACAAGCA
ACCCATGTTCTGCTGCAAGCTTGAAGGAGCCTGGAGCGGGAGAAAGCTAACTTGAACATGAC
CTGTTGCATTTGGCAAGTTCTAGCAACATGCTCCTAAGGAAGCGATACAGGCACAGACCATG
CAGACTCCAGTTCCTCCTGCTGCTCCTGATGCTGGGATGCGTCCTGATGATGGTGGCGATGT
TGCACCCTCCCCACCACACCCTGCACCAGACTGTCACAGCCCAAGCCAGCAAGCACAGCCCT
GAAGCCAGGTACCGCCTGGACTTTGGGGAATCCCAGGATTGGGTACTGGAAGCTGAGGATGA
GGGTGAAGAGTACAGCCCTCTGGAGGGCCTGCCACCCTTTATCTCACTGCGGGAGGATCAGC
TGCTGGTGGCCGTGGCCTTACCCCAGGCCAGAAGGAACCAGAGCCAGGGCAGGAGAGGTGGG
AGCTACCGCCTCATCAAGCAGCCAAGGAGGCAGGATAAGGAAGCCCCAAAGAGGGACTGGGG
GGCTGATGAGGACGGGGAGGTGTCTGAAGAAGAGGAGTTGACCCCGTTCAGCCTGGACCCAC
GTGGCCTCCAGGAGGCACTCAGTGCCCGCATCCCCCTCCAGAGGGCTCTGCCCGAGGTGCGG
CACCCACTGTGTCTGCAGCAGCACCCTCAGGACAGCCTGCCCACAGCCAGCGTCATCCTCTG
TTTCCATGATGAGGCCTGGTCCACTCTCCTGCGGACTGTACACAGCATCCTCGACACAGTGC
CCAGGGCCTTCCTGAAGGAGATCATCCTCGTGGACGACCTCAGCCAGCAAGGACAACTCAAG
TCTGCTCTCAGCGAATATGTGGCCAGGCTGGAGGGGGTGAAGTTACTCAGGAGCAACAAGAG
GCTGGGTGCCATCAGGGCCCGGATGCTGGGGGCCACCAGAGCCACCGGGGATGTGCTCGTCT
TCATGGATGCCCACTGCGAGTGCCACCCAGGCTGGCTGGAGCCCCTCCTCAGCAGAATAGCT
GGTGACAGGAGCCGAGTGGTATCTCCGGTGATAGATGTGATTGACTGGAAGACTTTCCAGTA
TTACCCCTCAAAGGACCTGCAGCGTGGGGTGTTGGACTGGAAGCTGGATTTCCACTGGGAAC
CTTTGCCAGAGCATGTGAGGAAGGCCCTCCAGTCCCCCATAAGCCCCATCAGGAGCCCTGTG
GTGCCCGGAGAGGTGGTGGCCATGGACAGACATTACTTCCAAAACACTGGAGCGTATGACTC
TCTTATGTCGCTGCGAGGTGGTGAAAACCTCGAACTGTCTTTCAAGGCCTGGCTCTGTGGTG
GCTCTGTTGAAATCCTTCCCTGCTCTCGGGTAGGACACATCTACCAAAATCAGGATTCCCAT
TCCCCCCTCGACCAGGAGGCCACCCTGAGGAACAGGGTTCGCATTGCTGAGACCTGGCTGGG
GTCATTCAAAGAAACCTTCTACAAGCATAGCCCAGAGGCCTTCTCCTTGAGCAAGGCTGAGA
AGCCAGACTGCATGGAACGCTTGCAGCTGCAAAGGAGACTGGGTTGTCGGACATTCCACTGG
TTTCTGGCTAATGTCTACCCTGAGCTGTACCCATCTGAACCCAGGCCCAGTTTCTCTGGAAA
GCTCCACAACACTGGACTTGGGCTCTGTGCAGACTGCCAGGCAGAAGGGGACATCCTGGGCT
GTCCCATGGTGTTGGCTCCTTGCAGTGACAGCCGGCAGCAACAGTACCTGCAGCACACCAGC
AGGAAGGAGATTCACTTTGGCAGCCCACAGCACCTGTGCTTTGCTGTCAGGCAGGAGCAGGT
GATTCTTCAGAACTGCACGGAGGAAGGCCTGGCCATCCACCAGCAGCACTGGGACTTCCAGG
AGAATGGGATGATTGTCCACATTCTTTCTGGGAAATGCATGGAAGCTGTGGTGCAAGAAAAC
AATAAAGATTTGTACCTGCGTCCGTGTGATGGAAAAGCCCGCCAGCAGTGGCGATTTGACCA
GATAAATGCTGTGGATGAACGATGAATGTCAATGTCAGAAGGAAAAGAGAATTTTGGCCATC
AAAATCCAGCTCCAAGTGAACGTAAAGAGCTTATATATTTCATGAAGCTGATCCTTTTGTGT
GTGTGCTCCTTGTGTTAGGAGAGAAAAAAGCTCTATGAAAGAATATAGGAAGTTTCTCCTTT
TCACACCTTATTTCATTGACTGCTGGCTGCTTA
```

FIGURE 202

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73760
><subunit 1 of 1, 639 aa, 1 stop
><MW: 73063, pI: 6.84, NX(S/T): 2

MLLRKRYRHRPCRLQFLLLLLMLGCVLMMVAMLHPPHHTLHQTVTAQASKHSPEARYRLDFG
ESQDWVLEAEDEGEEYSPLEGLPPFISLREDQLLVAVALPQARRNQSQGRRGGSYRLIKQPR
RQDKEAPKRDWGADEDGEVSEEEELTPFSLDPRGLQEALSARIPLQRALPEVRHPLCLQQHP
QDSLPTASVILCFHDEAWSTLLRTVHSILDTVPRAFLKEIILVDDLSQQGQLKSALSEYVAR
LEGVKLLRSNKRLGAIRARMLGATRATGDVLVFMDAHCECHPGWLEPLLSRIAGDRSRVVSP
VIDVIDWKTFQYYPSKDLQRGVLDWKLDFHWEPLPEHVRKALQSPISPIRSPVVPGEVVAMD
RHYFQNTGAYDSLMSLRGGENLELSFKAWLCGGSVEILPCSRVGHIYQNQDSHSPLDQEATL
RNRVRIAETWLGSFKETFYKHSPEAFSLSKAEKPDCMERLQLQRRLGCRTFHWFLANVYPEL
YPSEPRPSFSGKLHNTGLGLCADCQAEGDILGCPMVLAPCSDSRQQQYLQHTSRKEIHFGSP
QHLCFAVRQEQVILQNCTEEGLAIHQQHWDFQENGMIVHILSGKCMEAVVQENNKDLYLRPC
DGKARQQWRFDQINAVDER

Signal peptide:
amino acids 1-28

FIGURE 203

CGCCAAGCATGCAGTAAAGGCTGAAAATCTGGGTCACAGCTGAGGAAGACCTCAGAC<u>ATG</u>GA
GTCCAGGATGTGGCCTGCGCTGCTGCTGTCCCACCTCCTCCCTCTCTGGCCACTGCTGTTGC
TGCCCCTCCCACCGCCTGCTCAGGGCTCTTCATCCTCCCCTCGAACCCCACCAGCCCCAGCC
CGCCCCCGTGTGCCAGGGGAGGCCCCTCGGCCCCACGTCATGTGTGCGTGTGGGAGCGAGC
ACCTCCACCAAGCCGATCTCCTCGGGTCCCAAGATCACGTCGGCAAGTCCTGCCTGGCACTG
CACCCCCAGCCACCCCATCAGGCTTTGAGGAGGGGCCGCCCTCATCCCAATACCCCTGGGCT
ATCGTGTGGGGTCCCACCGTGTCTCGAGAGGATGGAGGGGACCCCAACTCTGCCAATCCCGG
ATTTCTGGACTATGGTTTTGCAGCCCCTCATGGGCTCGCAACCCCACACCCCAACTCAGACT
CCATGCGAGGTGATGGAGATGGGCTTATCCTTGGAGAGGCACCTGCCACCCTGCGGCCATTC
CTGTTCGGGGGCCGTGGGGAAGGTGTGGACCCCCAGCTCTATGTCACAATTACCATCTCCAT
CATCATTGTTCTCGTGGCCACTGGCATCATCTTCAAGTTCTGCTGGGACCGCAGCCAGAAGC
GACGCAGACCCTCAGGGCAGCAAGGTGCCCTGAGGCAGGAGGAGAGCCAGCAGCCACTGACA
GACCTGTCCCCGGCTGGAGTCACTGTGCTGGGGGCCTTCGGGGACTCACCTACCCCCACCCC
TGACCATGAGGAGCCCCGAGGGGGACCCCGGCCTGGGATGCCCCACCCCAAGGGGGCTCCAG
CCTTCCAGTTGAACCGG<u>TGA</u>GGGCAGGGGCAATGGGATGGGAGGGCAAAGAGGGAAGGCAAC
TTAGGTCTTCAGAGCTGGGGTGGGGGTGCCCTCTGGATGGGTAGTGAGGAGGCAGGCGTGGC
CTCCCACAGCCCCTGGCCCTCCCAAGGGGGCTGGACCAGCTCCTCTCTGGGAGGCACCCTTC
CTTCTCCCAGTCTCTCAGGATCTGTGTCCTATTCTCTGCTGCCCATAACTCCAACTCTGCCC
TCTTTGGTTTTTTCTCATGCCACCTTGTCTAAGACAACTCTGCCCTCTTAACCTTGATTCCC
CCTCTTTGTCTTGAACTTCCCCTTCTATTCTGGCCTACCCCTTGGTTCCTGACTGTGCCCTT
TCCCTCTTCCTCTCAGGATTCCCCTGGTGAATCTGTGATGCCCCAATGTTGGGGTGCAGCC
AAGCAGGAGGCCAAGGGGCCGGCACAGCCCCCATCCCACTGAGGGTGGGGCAGCTGTGGGGA
GCTGGGGCCACAGGGGCTCCTGGCTCCTGCCCCTTGCACACCACCCGGAACACTCCCCAGCC
CCACGGGCAATCCTATCTGCTCGCCCTCCTGCAGGTGGGGGCCTCACATATCTGTGACTTCG
GGTCCCTGTCCCCACCCTTGTGCACTCACATGAAAGCCTTGCACACTCACCTCCACCTTCAC
AGGCCATTTGCACACGCTCCTGCACCCTCTCCCCGTCCATACCGCTCCGCTCAGCTGACTCT
CATGTTCTCTCGTCTCACATTTGCACTCTCTCCTTCCCACATTCTGTGCTCAGCTCACTCAG
TGGTCAGCGTTTCCTGCACACTTTACCTCTCATGTGCGTTTCCCGGCCTGATGTTGTGGTGG
TGTGCGGCGTGCTCACTCTCTCCCTCATGAACACCCACCCACCTCGTTTCCGCAGCCCCTGC
GTGCTGCTCCAGAGGTGGGTGGGAGGTGAGCTGGGGGCTCCTTGGGCCCTCATCGGTCATGG
TCTCGTCCCATTCCACACCATTTGTTTCTCTGTCTCCCCATCCTACTCCAAGGATGCCGGCA
TCACCCTGAGGGCTCCCCCTTGGGAATGGGGTAGTGAGGCCCCAGACTTCACCCCCAGCCCA
CTGCTAAAATCTGTTTTCTGACAGATGGGTTTTGGGGAGTCGCCTGCTGCACTACATGAGAA
AGGGACTCCCATTTGCCCTTCCCTTTCTCCTACAGTCCCTTTTGTCTTGTCTGTCCTGGCTG
TCTGTGTGTGTGCCATTCTCTGGACTTCAGAGCCCCTGAGCCAGTCCTCCCTTCCCAGCCT
CCCTTTGGGCCTCCCTAACTCCACCTAGGCTGCCAGGGACCGGAGTCAGCTGGTTCAAGGCC
ATCGGGAGCTCTGCCTCCAAGTCTACCCTTCCCTTCCCGGACTCCCTCCTGTCCCCTCCTTT
CCTCCCTCCTTCCTTCCACTCTCCTTCCTTTTGCTTCCCTGCCCTTTCCCCCTCCTCAGGTT
CTTCCCTCCTTCTCACTGGTTTTTCCACCTTCCTCCTTCCCTTCTTCCCTGGCTCCTAGGCT
GTGATATATATTTTTGTATTATCTCTTTCTTCTTCTTGTGGTGATCATCTTGAATTACTGTG
GGATGTAAGTTTCAAAATTTTCAAATAAAGCCTTTGCAAGATAA

206/270

FIGURE 204

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393
><subunit 1 of 1, 243 aa, 1 stop
><MW: 26266, pI: 8.43, NX(S/T): 1
MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGV
PGRDGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA
ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIIYLDQGSPEMN
STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIEELPK
```

Signal peptide:
amino acids 1-30

Transmembrane domain:
amino acids 195-217

FIGURE 205

GTTAACCAGCGCAGTCCTCCGTGCGTCCCGCCCGCCGCTGCCCTCACTCCCGGCCAGG<u>ATG</u>G

CATCCTGTCTGGCCCTGCGCATGGCGCTGCTGCTGGTCTCCGGGGTTCTGGCCCCTGCGGTG

CTCACAGACGATGTTCCACAGGAGCCCGTGCCCACGCTGTGGAACGAGCCGGCCGAGCTGCC

GTCGGGAGAAGGCCCCGTGGAGAGCACCAGCCCCGGCCGGGAGCCCGTGGACACCGGTCCCC

CAGCCCCCACCGTCGCGCCAGGACCCGAGGACAGCACCGCGCAGGAGCGGCTGGACCAGGGC

GGCGGGTCGCTGGGGCCCGGCGCTATCGCGGCCATCGTGATCGCCGCCCTGCTGGCCACCTG

CGTGGTGCTGGCGCTCGTGGTCGTCGCGC<u>T</u>GAGAAAGTTTTCTGCCTCC<u>TGA</u>AGCGAATAAA

GGGGCCGCGCCCGGCCGCGGCGCGACTCGGCAAAAAAAAAAAAAAA

FIGURE 206

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76398

><subunit 1 of 1, 121 aa, 1 stop

><MW: 12073, pI: 4.11, NX(S/T): 0

MASCLALRMALLLVSGVLAPAVLTDDVPQEPVPTLWNEPAELPSGEGPVESTSPGREPVDTG

PPAPTVAPGPEDSTAQERLDQGGGSLGPGAIAAIVIAALLATCVVLALVVVALRKFSAS

Important features of the protein:
Signal peptide:
amino acids 1-19

Transmembrane domain:
amino acids 91-110

Glycosaminoglycan attachment site.
amino acids 44-47

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 116-119

N-myristoylation site.
amino acids 91-96

FIGURE 207

GGCCGTTGGTTGGTGCGCGGCTGAAGGGTGTGGCGCGAGCAGCGTCGTTGGTTGGCCGGCGG
CGGGCCGGGACGGGC**ATG**GCCCTGCTGCTGTGCCTGGTGTGCCTGACGGCGGCGCTGGCCCA
CGGCTGTCTGCACTGCCACAGCAACTTCTCCAAGAAGTTCTCCTTCTACCGCCACCATGTGA
ACTT̲CAAGTCCTGGTGGGTGGGCGACATCCCCGTGTCAGGGGCGCTGCTCACCGACTGGAGC
GACGACACGATGAAGGAGCTGCACCTGGCCATCCCCGCCAAGATCACCCGGGGAGAAGCTGGA
CCAAGTGGCGACAGCAGTGTACCAGATGATGGATCAGCTGTACCAGGGGAAGATGTACTTCC
CCGGGTATTTCCCCAACGAGCTGCGAAACATCTTCCGGGAGCAGGTGCACCTCATCCAGAAC
GCCATCATCGAAAGGCACCTGGCACCAGGCAGCTGGGGAGGAGGGCAGCTCTCCAGGGAGGG
ACCCAGCCTAGCACCTGAAGGATCAATGCCATCACCCCGCGGGGACCTCCCC**TAA**GTAGCCC
CCAGAGGCGCTGGGAGTGTTGCCACCGCCCTCCCCTGAAGTTTGCTCCATCTCACGCTGGGG
GTCAACCTGGGGACCCCTTCCCTCCGGGCCATGGACACACATACATGAAAACCAGGCCGCAT
CGACTGTCAGCACCGCTGTGGCATCTTCCAGTACGAGACCATCTCCTGCAACAACTGCACAG
ACTCGCACGTCGCCTGCTTTGGCTATAACTGCGAGTAGGGCTCAGGCATCACACCCACCCGT
GCCAGGGCCCTACTGTCCCTGGGGTCCCAGGCTCTCCTTGGAGGGGGCTCCCCGCCTTCCAC
CTGGCTGTCATCGGGTAGGGCGGGGCCGTGGGTTCAGGGGCGCACCCACTTCCAAGCCTGTGT
CCCACAGGTCCTCGGCGCAGTGGAAGTCAGCTGTCCAGGGCCTCCTGAACTACATAAATAAC
TGGCACAAGTAAGTCCCCTCCTCAAACCAACACAGGCAGTGTGTGTATGTGAGCACCTCGTG
GGTGAGTATGTGTGGGGCACAGGCTGGCTCCCTCAGCTCCCACGTCCTAGAGGGGCTCCCGA
GGAGGTGGAACCTCAACCCAGCTCTGCGCAGGAGGCGGCTGCAGTCCTTTTCTCCCTCAAAG
GTCTCCGACCCTCAGCTGGAGGCGGGCATCTTTCCTAAAGGGTCCCCATAGGGTCTGGTTCC
ACCCCATCCCAGGTCTGTGGTCAGAGCCTGGGAGGGTTCCCTACGATGGTTAGGGGTGCCCC
ATGGAGGGGCTGACTGCCCCACATTGCCTTTCAGACAGGACACGAGCATGAGGTAAGGCCGC
CCTGACCTGGACTTCAGGGGGAGGGGGTAAAGGGAGAGAGGAGGGGGGCTAGGGGGTCCTCT
AGATCAGTGGGGGCACTGCAGGTGGGGCTCTCCCTATACCTGGGACACCTGCTGGATGTCAC
CTCTGCAACCACACCCATGTGGTGGTTTCATGAACAGACCACGCTCCTCTGCCTTCTCCTGG
CCTGGGACACACAGAGCCACCCCGGCCTTGTGAGTGACCCAGAGAAGGGAGGCCTCGGGAGA
AGGGGTGCTCGTAAGCCAACACCAGCGTGCCGCGGCCTGCACACCCTTCGGACATCCCAGGC
ACGAGGGTGTCGTGGATGTGGCCACACATAGGACCACACGTCCCAGCTGGGAGGAGAGGCCT
GGGGCCCCCAGGGAGGGAGGCAGGGGGTGGGGGACATGGAGAGCTGAGGCAGCCTCGTCTCC
CCGCAGCCTGGTATCGCCAGCCTTAAGGTGTCTGGAGCCCCCACACTTGGCCAACCTGACCT
TGGAAGATGCTGCTGAGTGTCTCAAGCAGCACTGACAGCAGCTGGGCCTGCCCCAGGGCAAC
GTGGGGGCGGAGACTCAGCTGGACAGCCCCTGCCTGTCACTCTGGAGCTGGGCTGCTGCTGC
CTCAGGACCCCCTCTCCGACCCCGGACAGAGCTGAGCTGGCCAGGGCCAGGAGGGCGGGAGG
GAGGGAATGGGGGTGGGCTGTGCGCAGCATCAGCGCCTGGGCAGGTCCGCAGAGCTGCGGGA
TGTGATTAAAGTCCCTGATGTTTCTC

FIGURE 208

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76399

><subunit 1 of 1, 157 aa, 1 stop

><MW: 17681, pI: 7.65, NX(S/T): 1

MALLLCLVCLTAALAHGCLHCHSNFSKKFSFYRHHVNFKSWWVGDIPVSGALLTDWSDDTMK

ELHLAIPAKITREKLDQVATAVYQMMDQLYQGKMYFPGYFPNELRNIFREQVHLIQNAIIER

HLAPGSWGGGQLSREGPSLAPEGSMPSPRGDLP

Signal peptide:

amino acids 1-15

FIGURE 209

AGCAGGAGCAGGAGAGGGACA**ATG**GAAGCTGCCCCGTCCAGGTTCATGTTCCTCTTATTTCT

CCTCACGTGTGAGCTGGCTGCAGAAGTTGCTGCAGAAGTTGAGAAATCCTCAGATGGTCCTG

GTGCTGCCCAGGAACCCACGTGGCTCACAGATGTCCCAGCTGCCATGGAATTCATTGCTGCC

ACTGAGGTGGCTGTCATAGGCTTCTTCCAGGATTTAGAAATACCAGCAGTGCCCATACTCCA

TAGCATGGTGCAAAAATTCCCAGGCGTGTCATTTGGGATCAGCACTGATTCTGAGGTTCTGA

CACACTACAACATCACTGGGAACACCATCTGCCTCTTTCGCCTGGTAGACAATGAACAACTG

AATTTAGAGGACGAAGACATTGAAAGCATTGATGCCACCAAATTGAGCCGTTTCATTGAGAT

CAACAGCCTCCACATGGTGACAGAGTACAACCCTGTGACTGTGATTGGGTTATTCAACAGCG

TAATTCAGATTCATCTCCTCCTGATAATGAACAAGGCCTCCCCAGAGTATGAAGAGAACATG

CACAGATACCAGAAGGCAGCCAAGCTCTTCCAGGGGAAGATTCTCTTTATTCTGGTGGACAG

TGGTATGAAAGAAAATGGGAAGGTGATATCATTTTTTCAAACTAAAGGAGTCTCAACTGCCAG

CTTTGGCAATTTACCAGACTCTAGATGACGAGTGGGATACACTGCCCACAGCAGAAGTTTCC

GTAGAGCATGTGCAAAACTTTTGTGATGGATTCCTAAGTGGAAAATTGTTGAAAGAAAATCG

TGAATCAGAAGGAAAGACTCCAAAGGTGGAACTC**TGA**CTTCTCCTTGGAACTACATATGGCC

AAGTATCTACTTTATGCAAAGTAAAAAGGCACAACTCAAATCTCAGAGACACTAAACAACAG

GATCACTAGGCCTGCCAACCACACACACGCACGTGCACACACGCACGCACGCGTGCACAC

ACACACGCGCACACACACACACACAGAGCTTCATTTCCTGTCTTAAAATCTCGTTTTCTC

TTCTTCCTTCTTTTAAATTTCATATCCTCACTCCCTATCCAATTTCCTTCTTATCGTGCATT

CATACTCTGTAAGCCCATCTGTAACACACCTAGATCAAGGCTTTAAGAGACTCACTGTGATG

CCTCTATGAAAGAGAGGCATTCCTAGAGAAAGATTGTTCCAATTTGTCATTTAATATCAAGT

TTGTATACTGCACATGACTTACACACAACATAGTTCCTGCTCTTTTAAGGTTACCTAAGGGT

TGAAACTCTACCTTCTTTCATAAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCAAAGG

ATGGTTTTAAACACCTTTGTGAAATTGTCTTTTTGCCAGAAGTTAAAGGCTGTCTCCAAGTC

CCTGAACTCAGCAGAAATAGACCATGTGAAAACTCCATGCTTGGTTAGCATCTCCAACTCCC

TATGTAAATCAACAACCTGCATAATAAATAAAAGGCAATCATGTTATA

FIGURE 210

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76401

><subunit 1 of 1, 273 aa, 1 stop

><MW: 30480, pI: 4.60, NX(S/T): 1

MEAAPSRFMFLLFLLTCELAAEVAAEVEKSSDGPGAAQEPTWLTDVPAAMEFIAATEVAVIG

FFQDLEIPAVPILHSMVQKFPGVSFGISTDSEVLTHYNITGNTICLFRLVDNEQLNLEDEDI

ESIDATKLSRFIEINSLHMVTEYNPVTVIGLFNSVIQIHLLLIMNKASPEYEENMHRYQKAA

KLFQGKILFILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDEWDTLPTAEVSVEHVQNF

CDGFLSGKLLKENRESEGKTPKVEL

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 143-162

213/270

FIGURE 211

GGAGAGCCGCGGCTGGGACCGGAGTGGGGAGCGCGGCGTGGAGGTGCCACCCGGCGCGGGTG
GCGGAGAGATCAGAAGCCTCTTCCCCAAGCCGAGCCAACCTCAGCGGGGACCCGGGCTCAGG
GACGCGGCGGCGGCGGCGGCGACTGCAGTGGCTGGACG**ATG**GCAGCGTCCGCCGGAGCCGGG
GCGGTGATTGCAGCCCCAGACAGCCGGCGCTGGCTGTGGTCGGTGCTGGCGGCGGCGCTTGG
GCTCTTGACAGCTGGAGTATCAGCCTTGGAAGTATATACGCCAAAAGAAATCTTCGTGGCAA
ATGGTACACAAGGGAAGCTGACCTGCAAGTTCAAGTCTACTAGTACGACTGGCGGGTTGACC
TCAGTCTCCTGGAGCTTCCAGCCAGAGGGGGCCGACACTACTGTGTCGTTTTTCCACTACTC
CCAAGGGCAAGTGTACCTTGGGAATTATCCACCATTTAAAGACAGAATCAGCTGGGCTGGAG
ACCTTGACAAGAAAGATGCATCAATCAACATAGAAAATATGCAGTTTATACACAATGGCACC
TATATCTGTGATGTCAAAAACCCTCCTGACATCGTTGTCCAGCCTGGCACACATTAGGCTCTA
TGTCGTAGAAAAAGAGAATTTGCCTGTGTTTCCAGTTTGGGTAGTGGTGGGCATAGTTACTG
CTGTGGTCCTAGGTCTCACTCTGCTCATCAGCATGATTCTGGCTGTCCTCTATAGAAGGAAA
AACTCTAAACGGGATTACACTGGCTGCAGTACATCAGAGAGTTTGTCACCAGTTAAGCAGGC
TCCTCGGAAGTCCCCCTCCGACACTGAGGGTCTTGTAAAGAGTCTGCCTTCTGGATCTCACC
AGGGCCCAGTCATATATGCACAGTTAGACCACTCCGGCGGACATCACAGTGACAAGATTAAC
AAGTCAGAGTCTGTGGTGTATGCGGATATCCGAAAGAAT**TAA**GAGAATACCTAGAACATATC
CTCAGCAAGAAACAAAACCAAACTGGACTCTCGTGCAGAAAATGTAGCCCATTACCACATGT
AGCCTTGGAGACCCAGGCAAGGACAAGTACACGTGTACTCACAGAGGGAGAGAAAGATGTGT
ACAAAGGATATGTATAAATATTCTATTTAGTCATCCTGATATGAGGAGCCAGTGTTGCATGA
TGAAAGATGGTATGATTCTACATATGTACCCATTGTCTTGCTGTTTTTGTACTTTCTTTTC
AGGTCATTTACAATTGGGAGATTTCAGAAACATTCCTTTCACCATCATTTAGAAATGGTTTG
CCTTAATGGAGACAATAGCAGATCCTGTAGTATTTCCAGTAGACATGGCCTTTTAATCTAAG
GGCTTAAGACTGATTAGTCTTAGCATTTACTGTAGTTGGAGGATGGAGATGCTATGATGGAA
GCATACCCAGGGTGGCCTTTAGCACAGTATCAGTACCATTTATTTGTCTGCCGCTTTTAAAA
AATACCCATTGGCTATGCCACTTGAAAACAATTTGAGAAGTTTTTTTGAAGTTTTTCTCACT
AAAAATATGGGGCAATTGTTAGCCTTACATGTTGTGTAGACTTACTTTAAGTTTGCACCCTTG
AAATGTGTCATATCAATTTCTGGATTCATAATAGCAAGATTAGCAAAGGATAAATGCCGAAG
GTCACTTCATTCTGGACACAGTTGGATCAATACTGATTAAGTAGAAAATCCAAGCTTTGCTT
GAGAACTTTTGTAACGTGGAGAGTAAAAAGTATCGGTTTTA

FIGURE 212

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76510
><subunit 1 of 1, 269 aa, 1 stop
><MW: 29082, pI: 9.02, NX(S/T): 3
MAASAGAGAVIAAPDSRRWLWSVLAAALGLLTAGVSALEVYTPKEIFVANGTQGKLTCKFKS
TSTTGGLTSVSWSFQPEGADTTVSFFHYSQGQVYLGNYPPFKDRISWAGDLDKKDASINIEN
MQFIHNGTYICDVKNPPDIVVQPGHIRLYVVEKENLPVFPVWVVVGIVTAVVLGLTLLISMI
LAVLYRRKNSKRDYTGCSTSESLSPVKQAPRKSPSDTEGLVKSLPSGSHQGPVIYAQLDHSG
GHHSDKINKSESVVYADIRKN

Signal peptide:
amino acids 1-37

Transmembrane domain:
amino acids 161-183

FIGURE 213

GCCGGCTGTGCAGAGACGCC**ATG**TACCGGCTCCTGTCAGCAGTGACTGCCCGGGCTGCCGCC

CCCGGGGGGCTTGGCCTCAAGCTGCGGACGACGCGGGGTCCATCAGCGCGCCGGGCTGCCGCC

TCTCGGCCACGGCTGGGTCGGGGGCCTCGGGCTGGGGCTGGGGCTGGCGCTCGGGGTGAAGC

TGGCAGGTGGGCTGAGGGGCGCGGCCCCGGCGCAGTCCCCGCGGCCCCCGACCCTGAGGCG

TCGCCTCTGGCCGAGCCGCCACAGGAGCAGTCCCTCGCCCCGTGGTCTCCGCAGACCCCGGC

GCCGCCCTGCTCCAGGTGCTTCGCCAGAGCCATCGAGAGCAGCCGCGACCTGCTGCACAGGA

TCAAGGATGAGGTGGGCGCACCGGGCATÀGTGGTTGGAGTTTCTGTAGATGGAAAAGAAGTC

TGGTCAGAAGGTTTAGGTTATGCTGATGTTGAGAACCGTGTACCATGTAAACCAGAGACAGT

TATGCGAATTGCTAGCATCAGCAAAAGTCTCACCATGGTTGCTCTTGCCAAATTGTGGGAAG

CAGGGAAACTGGATCTTGATATTCCAGTACAACATTATGTTCCCGAATTCCCAGAAAAAGAA

TATGAAGGTGAAAAGGTTTCTGTCACAACAAGATTACTGATTTCCCATTTAAGTGGAATTCG

TCATTATGAAAAGGACATAAAAAAAGGTGAAAGAAGAGAAAGCTTATAAAGCCTTGAAGATGA

TGAAAGAGAATGTTGCATTTGAGCAAGAAAAAGAAGGCAAAAGTAATGAAAAGAATGATTTT

ACTAAATTTAAAACAGAGCAGGAGAATGAAGCCAAATGCCGGAATTCAAAACCTGGCAAGAA

AAAGAATGATTTTGAACAAGGCGAATTATATTTGAGAGAAAAGTTTGAAAATTCAATTGAAT

CCCTAAGATTATTTAAAAATGATCCTTTGTTCTTCAAACCTGGTAGTCAGTTTTTGTATTCA

ACTTTTGGCTATACCCTACTGGCAGCCATAGTAGAGAGAGCTTCAGGATGTAAATATTTGGA

CTATATGCAGAAAATATTCCATGACTTGGATATGCTGACGACTGTGCAGGAAGAAAACGAGC

CAGTGATTTACAATAGAGCAAGG**TAA**ATGAATACCTTCTGCTGTGTCTAGCTATATCGCATC

TTAACACTATTTTATTAATTAAAAGTCAAATTTTCTTTGTTTCCATTCCAAAATCAACCTGC

CACATTTTGGGAGCTTTTCTACATGTCTGTTTTCTCATCTGTAAAGTGAAGGAAGTAAAACA

TGTTTATAAAGTAAAAAAA

FIGURE 214

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76522
><subunit 1 of 1, 373 aa, 1 stop
><MW: 41221, pI: 8.54, NX(S/T): 0

MYRLLSAVTARAAAPGGLASSCGRRGVHQRAGLPPLGHGWVGGLGLGLGLALGVKLAGGLRG
AAPAQSPAAPDPEASPLAEPPQEQSLAPWSPQTPAPPCSRCFARAIESSRDLLHRIKDEVGA
PGIVVGVSVDGKEVWSEGLGYADVENRVPCKPETVMRIASISKSLTMVALAKLWEAGKLDLD
IPVQHYVPEFPEKEYEGEKVSVTTRLLISHLSGIRHYEKDIKKVKEEKAYKALKMMKENVAF
EQEKEGKSNEKNDFTKFKTEQENEAKCRNSKPGKKKNDFEQGELYLREKFENSIESLRLFKN
DPLFFKPGSQFLYSTFGYTLLAAIVERASGCKYLDYMQKIFHDLDMLTTVQEENEPVIYNRAR

Signal peptide:
amino acids 1-19

Transmembrane domain:
amino acids 39-60

217/270

FIGURE 215

GTGACACTATAGAAGAGCTATGACGTCGCATGCACGCGTACGTAAGCTCGGAATTCGGCTCG

AGGCTGGTGGGAAGAAGCCGAG**ATG**GCGGCAGCCAGCGCTGGGGCAACCCGGCTGCTCCTGC

TCTTGCTGATGGCGGTAGCAGCGCCCAGTCGAGCCCGGGGCAGCGGCTGCCGGGCCGGGACT

GGTGCGCGAGGGGCTGGGGCGGAAGGTCGAGAGGGCGAGGCCTGTGGCACGGTGGGGCTGCT

GCTGGAGCACTCATTTGAGATCGATGACAGTGCCAACTTCCGGAAGCGGGGCTCACTGCTCT

GGAACCAGCAGGATGGTACCTTGTCCCTGTCACAGCGGCAGCTCAGCGAGGAGGAGCGGGGC

CGACTCCGGGATGTGGCAGCCCTGAATGGCCTGTACCGGGTCCGGATCCCAAGGCGACCCGG

GGCCCTGGATGGCCTGGAAGCTGGTGGCTATGTCTCCTCCTTTGTCCCTGCGTGCTCCCTGG

TGGAGTCGCACCTGTCGGACCAGCTGACCCTGCACGTGGATGTGGCCGGCAACGTGGTGGGC

GTGTCGGTGGTGACGCACCCCGGGGGCTGCCGGGGCCATGAGGTGGAGGACGTGGACCTGGA

GCTGTTCAACACCTCGGTGCAGCTGCAGCCGCCCACCACAGCCCCAGGCCCTGAGACGGCGG

CCTTCATTGAGCGCCTGGAGATGGAACAGGCCCAGAAGGCCAAGAACCCCCAGGAGCAGAAG

TCCTTCTTCGCCAAATACTGGATGTACATCATTCCCGTCGTCCTGTTCCTCATGATGTCAGG

AGCGCCAGACACCGGGGGCCAGGGTGGGGGTGGGGGTGGGGGTGGTGGTGGGGGTAGTGGCC

TTTGCTGTGTGCCACCCTCCCTG**TAA**GTCTATTTAAAAACATCGACGATACATTGAAATGTG

TGAACGTTTTGAAAAGCTACAGCTTCCAGCAGCCAAAAGCAACTGTTGTTTTGGCAAGACGG

TCCTGATGTACAAGCTTGATTGAAATTCACTGCTCACTTGATACGTTATTCAGAAACCCAAG

GAATGGCTGTCCCCATCCTCATGTGGCTGTGTGGAGCTCAGCTGTGTTGTGTGGCAGTTTAT

TAAACTGTCCCCCAGATCGACACGCAAAAAAAAA

FIGURE 216

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76529

><subunit 1 of 1, 269 aa, 1 stop

><MW: 28004, pI: 5.80, NX(S/T): 1

MAAASAGATRLLLLLLMAVAAPSRARGSGCRAGTGARGAGAEGREGEACGTVGLLLEHSFEI

DDSANFRKRGSLLWNQQDGTLSLSQRQLSEEERGRLRDVAALNGLYRVRIPRRPGALDGLEA

GGYVSSFVPACSLVESHLSDQLTLHVDVAGNVVGVSVVTHPGGCRGHEVEDVDLELFNTSVQ

LQPPTTAPGPETAAFIERLEMEQAQKAKNPQEQKSFFAKYWMYIIPVVLFLMMSGAPDTGGQ

GGGGGGGGGGGSGLCCVPPSL

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 226-243

FIGURE 217

GGAGCGCTGCTGGAACCCGAGCCGGAGCCGGAGCCACAGCGGGGAGGGTGGCCTGGCGGCCT

GGAGCCGGACGTGTCCGGGGCGTCCCCGCAGACCGGGGCAGCAGGTCGTCCGGGGGCCCACC

ATGCTGGTGACTGCCTACCTTGCTTTTGTAGGCCTCCTGGCCTCCTGCCTGGGGCTGGAACT

GTCAAGATGCCGGGCTAAACCCCCTGGAAGGGCCTGCAGCAATCCCTCCTTCCTTCGGTTTC

AACTGGACTTCTATCAGGTCTACTTCCTGGCCCTGGCAGCTGATTGGCTTCAGGCCCCCTAC

CTCTATAAACTCTACCAGCATTACTACTTCCTGGAAGGTCAAATTGCCATCCTCTATGTCTG

TGGCCTTGCCTCTACAGTCCTCTTTGGCCTAGTGGCCTCCTCCCTTGTGGATTGGCTGGGTC

GCAAGAATTCTTGTGTCCTCTTCTCCCTGACTTACTCACTATGCTGCTTAACCAAACTCTCT

CAAGACTACTTTGTGCTGCTAGTGGGGCGAGCACTTGGTGGGCTGTCCACAGCCCTGCTCTT

CTCAGCCTTCGAGGCCTGGTATATCCATGAGCACGTGGAACGGCATGACTTCCCTGCTGAGT

GGATCCCAGCTACCTTTGCTCGAGCTGCCTTCTGGAACCATGTGCTGGCTGTAGTGGCAGGT

GTGGCAGCTGAGGCTGTAGCCAGCTGGATAGGGCTGGGGCCTGTAGCGCCCTTTGTGGCTGC

CATCCCTCTCCTGGCTCTGGCAGGGGCCTTGGCCCTTCGAAACTGGGGGGAGAACTATGACC

GGCAGCGTGCCTTCTCAAGGACCTGTGCTGGAGGCCTGCGCTGCCTCCTGTCGGACCGCCGC

GTGCTGCTGCTGGGCACCATACAAGCTCTATTTGAGAGTGTCATCTTCATCTTTGTCTTCCT

CTGGACACCTGTGCTGGACCCACACGGGGCCCCTCTGGGCATTATCTTCTCCAGCTTCATGG

CAGCCAGCCTGCTTGGCTCTTCCCTGTACCGTATCGCCACCTCCAAGAGGTACCACCTTCAG

CCCATGCACCTGCTGTCCCTTGCTGTGCTCATCGTCGTCTTCTCTCTCTTCATGTTGACTTT

CTCTACCAGCCCAGGCCAGGAGAGTCCGGTGGAGTCCTTCATAGCCTTTCTACTTATTGAGT

TGGCTTGTGGATTATACTTTCCCAGCATGAGCTTCCTACGGAGAAAGGTGATCCCTGAGACA

GAGCAGGCTGGTGTACTCAACTGGTTCCGGGTACCTCTGCACTCACTGGCTTGCCTAGGGCT

CCTTGTCCTCCATGACAGTGATCGAAAAACAGGCACTCGGAATATGTTCAGCATTTGCTCTG

CTGTCATGGTGATGGCTCTGCTGGCAGTGGTGGGACTCTTCACCGTGGTAAGGCATGATGCT

GAGCTGCGGGTACCTTCACCTACTGAGGAGCCCTATGCCCCTGAGCTG**TAA**CCCCACTCCAG

GACAAGATAGCTGGGACAGACTCTTGAATTCCAGCTATCCGGGATTGTACAGATCTCTCTGT

GACTGACTTTGTGACTGTCCTGTGGTTTCTCCTGCCATTGCTTTGTGTTTGGGAGGACATGA

TGGGGGTGATGGACTGGAAAGAAGGTGCCAAAAGTTCCCTCTGTGTTACTCCCATTTAGAAA

ATAAACACTTTTAAATGATCAAAAAAAAAAAA

FIGURE 218

MLVTAYLAFVGLLASCLGLELSRCRAKPFGRACSNPSFLRFQLDFYQVYFLALAADWLQAPY

LYKLYQHYYFLEGQIAILYVCGLASTVLFGLVASSLVDWLGRKNSCVLFSLTYSLCCLTKLS

QDYFVLLVGRALGGLSTALLFSAFEAWYIHEHVERHDFPAEWIPATFARAAFWNHVLAVVAG

VAAEAVASWIGLGPVAPFVAAIPLLALAGALALRNWGENYDRQRAFSRTCAGGLRCLLSDRR

VLLLGTIQALFESVIFIFVFLWTPVLDPHGAPLGIIFSSFMAASLLGSSLYRIATSKRYHLQ

PMHLLSLAVLIVVFSLFMLTFSTSPGQESPVESFIAFLLIELACGLYFPSMSFLRRKVIPET

EQAGVLNWFRVPLHSLACLGLLVLHDSDRKTGTRNMFSICSAVMVMALLAVVGLFTVVRHDA

ELRVPSPTEEPYAPEL

 Signal peptide:
amino acids 1-18

Transmembrane domain:
amino acids 41-55, 75-94, 127-143, 191-213, 249-270, 278-299,
314-330, 343-359, 379-394, 410-430

FIGURE 219

GCGACGCGCGGCGGGGCGGCGAGAGGAAACGCGGCGCCGGGCCGGGCCCGGCCCTGGAG**ATG**

GTCCCCGGCGCCGCGGGCTGGTGTTGTCTCGTGCTCTGGCTCCCCGCGTGCGTCGCGGCCCA

CGGCTTCCGTATCCATGATTATTTGTACTTTCAAGTGCTGAGTCCTGGGGACATTCGATACA

TCTTCACAGCCACACCTGCCAAGGACTTTGGTGGTATCTTTCACACAAGGTATGAGCAGATT

CACCTTGTCCCCGCTGAACCTCCAGAGGCCTGCGGGGAACTCAGCAACGGTTTCTTCATCCA

GGACCAGATTGCTCTGGTGGAGAGGGGGGGCTGCTCCTTCCTCTCCAAGACTCGGGTGGTCC

AGGAGCACGGCGGGCGGGCGGTGATCATCTCTGACAACGCAGTTGACAATGACAGCTTCTAC

GTGGAGATGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCGCCCTCTTCCTGCTCGG

CCGAGACGGCTACATGATCCGCCGCTCTCTGGAACAGCATGGGCTGCCATGGGCCATCATTT

CCATCCCAGTCAATGTCACCAGCATCCCCACCTTTGAGCTGCTGCAACCGCCCTGGACCTTC

TGG**TAG**AAGAGTTTGTCCCACATTCCAGCCATAAGTGACTCTGAGCTGGGAAGGGGAAACCC

AGGAATTTTGCTACTTGGAATTTGGAGATAGCATCTGGGGACAAGTGGAGCCAGGTAGAGGA

AAAGGGTTTGGGCGTTGCTAGGCTGAAAGGGAAGCCACACCACTGGCCTTCCCTTCCCCAGG

GCCCCCAAGGGTGTCTCATGCTACAAGAAGAGGCAAGAGACAGGCCCCAGGGCTTCTGGCTA

GAACCCGAAACAAAAGGAGCTGAAGGCAGGTGGCCTGAGAGCCATCTGTGACCTGTCACACT

CACCTGGCTCCAGCCTCCCCTACCCAGGGTCTCTGCACAGTGACCTTCACAGCAGTTGTTGG

AGTGGTTTAAAGAGCTGGTGTTTGGGGACTCAATAAACCCTCACTGACTTTTTAGCAATAAA

GCTTCTCATCAGGGTTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 220

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76532

><subunit 1 of 1, 188 aa, 1 stop

><MW: 21042, pI: 5.36, NX(S/T): 2

MVPGAAGWCCLVLWLPACVAAHGFRIHDYLYFQVLSPGDIRYIFTATPAKDFGGIFHTRYEQ

IHLVPAEPPEACGELSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVIISDNAVDNDSF

YVEMIQDSTQRTADIPALFLLGRDGYMIRRSLEQHGLPWAIISIPVNVTSIPTFELLQPPWTFW

Signal peptide:

amino acids 1-20

FIGURE 221

TCTGCCTCCACTGCTCTGTGCTGGGATC**ATG**GAACTTGCACTGCTGTGTGGGCTGGTGGTGA

TGGCTGGTGTGATTCCAATCCAGGGCGGGATCCTGAACCTGAACAAGATGGTCAAGCAAGTG

ACTGGGAAAATGCCCATCCTCTCCTACTGGCCCTACGGCTGTCACTGCGGACTAGGTGGCAG

AGGCCAACCCAAAGATGCCACGGACTGGTGCTGCCAGACCCATGACTGCTGCTATGACCACC

TGAAGACCCAGGGGTGCGGCATCTACAAGGACAACAACAAAAGCAGCATACATTGTATGGAT

TTATCTCAACGCTATTGTTTAATGGCTGTGTTTAATGTGATCTATCTGGAAAATGAGGACTC

CGAA**TAA**AAAGCTATTACTAWTTNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 222

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76538

><subunit 1 of 1, 116 aa, 1 stop

><MW: 12910, pI: 6.41, NX(S/T): 1

MELALLCGLVVMAGVIPIQGGILNLNKMVKQVTGKMPILSYWPYGCHCGLGGRGQPKDATDW

CCQTHDCCYDHLKTQGCGIYKDNNKSSIHCMDLSQRYCLMAVFNVIYLENEDSE

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 1-24

N-glycosylation site.

amino acids 86-89

N-myristoylation sites.

amino acids 20-25, 45-50

Phospholipase A2 histidine active site.

amino acids 63-70

FIGURE 223

CTCGCTTCTTCCTTCTGGATGGGGGCCCAGGGGGCCCAGGAGAGTATAAAGGCGATGTGGAG

GGTGCCCGGCACAACCAGACGCCCAGTCACAGGCGAGAGCCCTGGG<u>ATG</u>CACCGGCCAGAGG

CCATGCTGCTGCTGCTCACGCTTGCCCTCCTGGGGGGCCCCACCTGGGCAGGGAAGATGTAT

GGCCCTGGAGGAGGCAAGTATTTCAGCACCACTGAAGACTACGACCATGAAATCACAGGGCT

GCGGGTGTCTGTAGGTCTTCTCCTGGTGAAAAGTGTCCAGGTGAAACTTGGAGACTCCTGGG

ACGTGAAACTGGGAGCCTTAGGTGGGAATACCCAGGAAGTCACCCTGCAGCCAGGCGAATAC

ATCACAAAAGTCTTTGTCGCCTTCCAAGCTTTCCTCCGGGGTATGGTCATGTACACCAGCAA

GGACCGCTATTTCTATTTTGGGAAGCTTGATGGCCAGATCTCCTCTGCCTACCCCAGCCAAG

AGGGGCAGGTGCTGGTGGGCATCTATGGCCAGTATCAACTCCTTGGCATCAAGAGCATTGGC

TTTGAATGGAATTATCCACTAGAGGAGCCGACCACTGAGCCACCAGTTAATCTCACATACTC

AGCAAACTCACCCGTGGGTCGC<u>TAG</u>GGTGGGGTATGGGGCCATCCGAGCTGAGGCCATCTGT

GTGGTGGTGGCTGATGGTACTGGAGTAACTGAGTCGGGACGCTGAATCTGAATCCACCAATA

AATAAAGCTTCTGCAGAAAA

FIGURE 224

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76541

><subunit 1 of 1, 178 aa, 1 stop

><MW: 19600, pI: 5.89, NX(S/T): 1

MHRPEAMLLLLTLALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSVGLLLVKSVQVK

LGDSWDVKLGALGGNTQEVTLQPGEYITKVFVAFQAFLRGMVMYTSKDRYFYFGKLDGQISS

AYPSQEGQVLVGIYGQYQLLGIKSIGFEWNYPLEEPTTEPPVNLTYSANSPVGR

Signal peptide:
amino acids 1-22

FIGURE 225

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCT
GAACTGGGTGCTCATCACGGGAACTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAG
CCCCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTT
TTTTTTTTAACCGCCCCCTCCCCACCCCCCAAAAAAACTGTAAAGATGCAAAAACGTAATAT
CCATGAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATT
TATTTGTTCTTGGAGTGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCT
CCCAAGGGGTCCAATTTTTCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTG
ACAGGGGCTGTCATGCAACTGGCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAA
CAATACAAAGG<u>ATG</u>GGTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTT
ATAGCCCCCACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTG
TAGGTGTGAAGGCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCCTCAAGTA
TATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAACTTAAGTATAAT
CAATTTAAAGGGCTCAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATAT
TGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCCAATAGAA
TCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATTTACGGAACTTGGATCTGTCC
TATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTGCGGAAGCTGCTGAGTTT
ACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACC
TGGAACTTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGC
ATGATCAGACTCAAAGAACTTCACCTGGAGCACAATCAATTTTCCAAGCTCAACCTGGCCCT
TTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGAATAAAATCAGTGTCATAG
GACAGACCATGTCCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATC
GAAGCTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGA
TTCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACA
TCAGTCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAACTGGCTG
AAAAGTTTTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGG
AGTAAATGTGATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGT
TTGATCTGGCCAGGGCTCTCCCAAAGCCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCAT
GAGAGCAAACCCCCTTTGCCCCCGACGGTGGGAGCCACAGAGCCCGGCCCAGAGACCGATGC
TGACGCCGAGCACATCTCTTTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCCTGTCCG
TGCTCGTCATCCTGCTGGTTATCTACGTGTCATGGAAGCGGTACCCTGCGAGCATGAAGCAG
CTGCAGCAGCGCTCCCTCATGCGAAGGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAAT
GACTCCCAGCACCCAGGAATTTTATGTAGATTATAAACCCACCAACACGGAGACCAGCGAGA
TGCTGCTGAATGGGACGGGACCCTGCACCTATAACAAATCGGGCTCCAGGGAGTGTGAGGTA
<u>TGA</u>ACCATTGTGATAAAAAGAGCTCTTAAAAGCTGGGAAATAAGTGGTGCTTTATTGAACTC
TGGTGACTATCAAGGGAACGCGATGCCCCCCCTCCCCTTCCCTCTCCCTCTCACTTTGGTGG
CAAGATCCTTCCTTGTCCGTTTTAGTGCATTCATAATACTGGTCATTTTCCTCTCATACATA
ATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTTGAACTCCGGTTTAATATAA
TACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTCGCATTTGTTTTAAGATAAAACT
TCTTTCATAGGTAAAAAAAAAAA

FIGURE 226

></usr/seqdb2/set/DNA/Dnaseqs.min/ss.DNA77301

><subunit 1 of 1, 513 aa, 1 stop

><MW: 58266, pI: 9.84, NX(S/T): 4

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAG
CLGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYF
LNNTFRPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELL
DLGYNRIRSLARNVFAGMIRLKELHLEHNQFSKLNLALFPRLVSLQNLYLQWNKISVIGQTM
SWTWSSLQRLDLSGNEIEAFSGPSVFQCVPNLQRLNLDSNKLTFIGQEILDSWISLNDISLA
GNIWECSRNICSLVNWLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLA
RALPKPTFKPKLPRPKHESKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVI
LLVIYVSWKRYPASMKQLQQRSLMRRHRKKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLN
GTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 227

AGTTCTGAGAAAGAAGGAAATAAACACAGGCACCAAACCACTATCCTAAGTTGACTGTCCTT

TAAATA<u>ATG</u>TCAAGATCCAGACTTTTCAGTGTCACCTCAGCGATCTCAACGATAGGGATCTTG

TGTTTGCCGCTATTCCAGTTGGTGCTCTCGGACCTACCATGCGAAGAAGATGAAATGTGTGT

AAATTATAATGACCAACACCCTAATGGCTGGTATATCTGGATCCTCCTGCTGCTGGTTTTGG

TGGCAGCTCTTCTCTGTGGAGCTGTGGTCCTCTGCCTCCAGTGCTGGCTGAGGAGACCCCGA

ATTGATTCTCACAGGCGCACCATGGCAGTTTTTGCTGTTGGAGACTTGGACTCTATTTATGG

GACAGAAGCAGCTGTGAGTCCAACTGTTGGAATTCACCTTCAAACTCAAACCCCTGACCTAT

ATCCTGTTCCTGCTCCATGTTTTGGCCCTTTAGGCTCCCCACCTCCATATGAAGAAATTGTA

AAAACAACC<u>TGA</u>TTTTAGGTGTGGATTATCAATTTAAAGTATTAACGACATCTGTAATTCCA

AAACATCAAATTTAGGAATAGTTATTTCAGTTGTTGGAAATGTCCAGAGATCTATTCATATA

GTCTGAGGAAGGACAATTCGACAAAAGAATGGATGTTGGAAAAAATTTTGGTCATGGAGATG

TTTAAATAGTAAAGTAGCAGGCTTTTGATGTGTCACTGCTGTATCATACTTTTATGCTACAC

AACCAAATTAATGCTTCTCCACTAGTATCCAAACAGGCAACAATTAGGTGCTGGAAGTAGTT

TCCATCACATTTAGGACTCCACTGCAGTATACAGCACACCATTTTCTGCTTTAAACTCTTTC

CTAGCATGGGGTCCATAAAAATTATTATAATTTAACAATAGCCCAAGCCGAGAATCCAACAT

GTCCAGAACCAGAACCAGAAAGATAGTATTTGAATGAAGGTGAGGGGAGAGAGTAGGAAAAA

GAAAAGTTTGGAGTTGAAGGGTAAAGGATAAATGAAGAGGAAAAGGAAAAGATTACAAGTCT

CAGCAAAAACAAGAGGTTTTATGCCCCAACCTGAAGAGGAAGAAATTGTAGATAGAAGGTGA

AGGAGATTGCTGAAGATATAGAGCACATATAATGCCAACACGGGGAGAAAAGAAAATTTCCC

CTTTTACAGTAATGAATGTGGCCTCCATAGTCCATAGTGTTTCTCTGGAGCCTCAGGGCTTG

GCATTTATTGCAGCATCATGCTAAGAACCTTCGGCATAGGTATCTGTTCCCATGAGGACTGC

AGAAGTAGCAATGAGACATCTTCAAGTGGCATTTTGGCAGTGGCCATCAGCAGGGGGACAGA

CAAAAACATCCATCACAGATGACATATGATCTTCAGCTGACAAATTTGTTGAACAAAACAAT

AAACATCAATAGATATCTAAAAA

FIGURE 228

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77303

><subunit 1 of 1, 146 aa, 1 stop

><MW: 16116, pI: 4.99, NX(S/T): 0

MSRSRLFSVTSAISTIGILCLPLFQLVLSDLPCEEDEMCVNYNDQHPNGWYIWILLLLVLVA

ALLCGAVVLCLQCWLRRPRIDSHRRTMAVFAVGDLDSIYGTEAAVSPTVGIHLQTQTPDLYP

VPAPCFGPLGSPPPYEEIVKTT

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 52-70

FIGURE 229

```
GAGCGGAGTAAAATCTCCACAAGCTGGGAACAAACCTCGTCCCAACTCCCACCCACCGGCGT
TTCTCCAGCTCGATCTGGAGGCTGCTTCGCCAGTGTGGGACGCAGCTGACGCCCGCTTATTA
GCTCTCGCTGCGTCGCCCCGGCTCAGAAGCTCCGTGGCGGCGGCGACCGTGACGAGAAGCCC
ACGGCCAGCTCAGTTCTCTTCTACTTTGGGAGAGAGAGAAAGTCAGATGCCCCTTTTAAACT
CCCTCTTCAAAACTCATCTCCTGGGTGACTGAGTTAATAGAGTGGATACAACCTTGCTGAAG
ATGAAGAATATACAATATTGAGGATATTTTTTTCTTTTTTTTTTCAAGTCTTGATTTGTGGC
TTACCTCAAGTTACCATTTTTCAGTCAAGTCTGTTTGTTTGCTTCTTCAGAAATGTTTTTTA
CAATCTCAAGAAAAAATATGTCCCAGAAATTGAGTTTACTGTTGCTTGTATTTGGACTCATT
TGGGGATTGATGTTACTGCACTATACTTTTCAACAACCAAGACATCAAAGCAGTGTCAAGTT
ACGTGAGCAAATACTAGACTTAAGCAAAAGATATGTTAAAGCTCTAGCAGAGGAAAATAAGA
ACACAGTGGATGTCGAGAACGGTGCTTCTATGGCAGGATATGCGGATCTGAAAAGAACAATT
GCTGTCCTTCTGGATGACATTTTGCAACGATTGGTGAAGCTGGAGAACAAAGTTGACTATAT
TGTTGTGAATGGCTCAGCAGCCAACACCACCAATGGTACTAGTGGGAATTTGGTGCCAGTAA
CCACAAATAAAAGAACGAATGTCTCGGGCAGTATCAGATAGCAGTTGAAAATCACCTTGTGC
TGCTCCATCCACTGTGGATTATATCCTATGGCAGAAAAGCTTTATAATTGCTGGCTTAGGAC
AGAGCAATACTTTACAATAAAAGCTCTACACATTTTCAAGGAGTATGCTGGATTCATGGAAC
TCTAATTCTGTACATAAAAATTTTAAAGTTATTTGTTTGCTTTCAGGCAAGTCTGTTCAATG
CTGTACTATGTCCTTAAAGAGAATTTGGTAACTTGGTTGATGTGGTAAGCAGATAGGTGAGT
TTTGTATAAATCTTTTGTGTTTGAGATCAAGCTGAAATGAAAACACTGAAAAACATGGATTC
ATTTCTATAACACATTTATTTAAGTATATAACACGTTTTTTGGACAAGTGAAGAATGTTTAA
TCATTCTGTCATTTGTTCTCAATAGATGTAACTGTTAGACTACGGCTATTTGAAAAAATGTG
CTTATTGTACTATATTTTGTTATTCCAATTATGAGCAGAGAAAGGAAATATAATGTTGAAAA
TAATGTTTTGAAATCATGACCCAAAGAATGTATTGATTTGCACTATCCTTCAGAATAACTGA
AGGTTAATTATTGTATATTTTTAAAAAATTACACTTATAAGAGTATAATCTTGAAATGGGTAG
CAGCCACTGTCCATTACCTATCGTAAACATTGGGGCAATTTAATAACAGCATTAAAATAGTT
GTAAACTCTAATCTTATACTTATTGAAGAATAAAAGATATTTTTATGATGAGAGTAACAATA
AAGTATTCATGATTTTTCACATACATGAATGTTCATTTAAAAGTTTAATCCTTTGAGTGTCT
ATGCTATCAGGAAAGCACATTATTTCCATATTTGGGTTAATTTTGCTTTTATTATATTGGTC
TAGGAGGAAGGGACTTTGGAGAATGGAACTCTTGAGGACTTTAGCCAGGTGTATATAATAAA
GGTACTTTTGTGCTGCATTAAATTGCTTGGAAAGTGTTAACATTATATTATATAAGAGTATC
CTTTATGAAATTTTGAATTTGTATAACAGATGCATTAGATATTCATTTTATATAATGGCCAC
TTAAAATAAGAACATTTAAAATATAAACTATGAAGATTGACTATCTTTTCAGGAAAAAAGCT
GTATATAGCACAGGGAACCCTAATCTTGGGTAATTCTAGTATAAAACAAATTATACTTTTAT
TTAAATTTCCCTTGTAGCAAATCTAATTGCCACATGGTGCCCTATATTTCATAGTATTTATT
CTCTATAGTAACTGCTTAAGTGCAGCTAGCTTCTAGATTTAGACTATATAGAATTTAGATAT
TGTATTGTTCGTCATTATAATATGCTACCACATGTAGCAATAATTACAATATTTTATTAAAA
TAAATATGTGAAATATTGTTTCATGAAAGACAGATTTCCAAATCTCTCTTCTCTTCTCTGTA
CTGTCTACCTTTATGTGAAGAAATTAATTATATGCCATTGCCAGGT
```

FIGURE 230

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77648

><subunit 1 of 1, 140 aa, 1 stop.

><MW: 15668, pI: 10.14, NX(S/T): 5

MFFTISRKNMSQKLSLLLLVFGLIWGLMLLHYTFQQPRHQSSVKLREQILDLSKRYVKALAE

ENKNTVDVENGASMAGYADLKRTIAVLLDDILQRLVKLENKVDYIVVNGSAANTTNGTSGNL

VPVTTNKRTNVSGSIR

Important features of the protein:

Signal peptide:

amino acids 1-26

233/270

FIGURE 231

CGCGGCCGGGCCGCCGGGGTGAGCGTGCCGAGGCGGCTGTGGCGCAGGCTTCCAGCCCCCAC
C<u>ATG</u>CCGTGGCCCCTGCTGCTGCTGCTGGCCGTGAGTGGGGCCCAGACAACCCGGCCATGCT
TCCCCGGGTGCCAATGCGAGGTGGAGACCTTCGGCCTTTTCGACAGCTTCAGCCTGACTCGG
GTGGATTGTAGCGGCCTGGGCCCCCACATCATGCCGGTGCCCATCCCTCTGGACACAGCCCA
CTTGGACCTGTCCTCCAACCGGCTGGAGATGGTGAATGAGTCGGTGTTGGCGGGGCCGGGCT
ACACGACGTTGGCTGGCCTGGATCTCAGCCACAACCTGCTCACCAGCATCTCACCCACTGCC
TTCTCCCGCCTTCGCTACCTGGAGTCGCTTGACCTCAGCCACAATGGCCTGACAGCCCTGCC
AGCCGAGAGCTTCACCAGCTCACCCCTGAGCGACGTGAACCTTAGCCACAACCAGCTCCGGG
AGGTCTCAGTGTCTGCCTTCACGACGCACAGTCAGGGCCGGGCACTACACGTGGACCTCTCC
CACAACCTCATTCACCGCCTCGTGCCCCACCCCACGAGGGCCGGCCTGCCTGCGCCCACCAT
TCAGAGCCTGAACCTGGCCTGGAACCGGCTCCATGCCGTGCCCAACCTCCGAGACTTGCCCC
TGCGCTACCTGAGCCTGGATGGGAACCCTCTAGCTGTCATTGGTCCGGGTGCCTTCGCGGGG
CTGGGAGGCCTTACACACCTGTCTCTGGCCAGCCTGCAGAGGCTCCCTGAGCTGGCGCCCAG
TGGCTTCCGTGAGCTACCGGGCCTGCAGGTCCTGGACCTGTCGGGCAACCCCAAGCTTAACT
GGGCAGGAGCTGAGGTGTTTTCAGGCCTGAGCTCCCTGCAGGAGCTGGACCTTTCGGGCACC
AACCTGGTGCCCCTGCCTGAGGCGCTGCTCCTCCACCTCCCGGCACTGCAGAGCGTCAGCGT
GGGCCAGGATGTGCGGTGCCGGCGCCTGGTGCGGGAGGGCACCTACCCCCGGAGGCCTGGCT
CCAGCCCCAAGGTGCCCCTGCACTGCGTAGACACCCGGGAATCTGCTGCCAGGGGCCCCACC
ATCTT<u>GTGA</u>CAAATGGTGTGGCCCAGGGCCACATAACAGACTGCTGTCCTGGGCTGCCTCAG
GTCCCGAGTAACTTATGTTCAATGTGCCAACACCAGTGGGGAGCCCGCAGGCCTATGTGGCA
GCGTCACCACAGGAGTTGTGGGCCTAGGAGAGGCTTTGGACCTGGGAGCCACACCTAGGAGC
AAAGTCTCACCCCTTTGTCTACGTTGCTTCCCCAAACCATGAGCAGAGGGACTTCGATGCCA
AACCAGACTCGGGTCCCCTCCTGCTTCCCTTCCCCACTTATCCCCCAAGTGCCTTCCCTCAT
GCCTGGGCCGGCCTGACCCGCAATGGGCAGAGGGTGGGTGGGACCCCCTGCTGCAGGGCAGA
GTTCAGGTCCACTGGGCTGAGTGTCCCCTTGGGCCCATGGCCCAGTCACTCAGGGGCGAGTT
TCTTTTCTAACATAGCCCTTTCTTTGCCATGAGGCCATGAGGCCCGCTTCATCCTTTTCTAT
TTCCCTAGAACCTTAATGGTAGAAGGAATTGCAAAGAATCAAGTCCACCCTTCTCATGTGAC
AGATGGGGAAACTGAGGCCTTGAGAAGGAAAAAGGCTAATCTAAGTTCCTGCGGGCAGTGGC
ATGACTGGAGCACAGCCTCCTGCCTCCCAGCCCGGACCCAATGCACTTTCTTGTCTCCTCTA
ATAAGCCCCACCCTCCCCGCCTGGGCTCCCCTTGCTGCCCTTGCCTGTTCCCCATTAGCACA
GGAGTAGCAGCAGCAGGACAGGCAAGAGCCTCACAAGTGGGACTCTGGGCCTCTGACCAGCT
GTGCGGCATGGGCTAAGTCACTCTGCCCTTCGGAGCCTCTGGAAGCTTAGGGCACATTGGTT
CCAGCCTAGCCAGTTTCTCACCCTGGGTTGGGGTCCCCCAGCATCCAGACTGGAAACCTACC
CATTTTCCCCTGAGCATCCTCTAGATGCTGCCCCAAGGAGTTGCTGCAGTTCTGGAGCCTCA
TCTGGCTGGGATCTCCAAGGGGCCTCCTGGATTCAGTCCCCACTGGCCCTGAGCACGACAGC
CCTTCTTACCCTCCCAGGAATGCCGTGAAAGGAGACAAGGTCTGCCCGACCCATGTCTATGC
TCTACCCCCAGGGCAGCATCTCAGCTTCCGAACCCTGGGCTGTTTCCTTAGTCTTCATTTTA
TAAAAGTTGTTGCCTTTTTAACGGAGTGTCACTTTCAACCGGCCTCCCCTACCCCTGCTGGC
CGGGGATGGAGACATGTCATTTGTAAAAGCAGAAAAAGGTTGCATTTGTTCACTTTTGTAAT
ATTGTCCTGGGCCTGTGTTGGGGTGTTGGGGGAAGCTGGGCATCAGTGGCCACATGGGCATC
AGGGGCTGGCCCCACAGAGACCCCACAGGGCAGTGAGCTCTGTCTTCCCCCACCTGCCTAGC
CCATCATCTATCTAACCGGTCCTTGATTTAATAAACACTATAAAAGGTTTAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

234/270

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77652

><subunit 1 of 1, 353 aa, 1 stop

><MW: 37847, pI: 6.80, NX(S/T): 2

MPWPLLLLLAVSGAQTTRPCFPGCQCEVETFGLFDSFSLTRVDCSGLGPHIMPVPIPLDTAH

LDLSSNRLEMVNESVLAGPGYTTLAGLDLSHNLLTSISPTAFSRLRYLESLDLSHNGLTALP

AESFTSSPLSDVNLSHNQLREVSVSAFTTHSQGRALHVDLSHNLIHRLVPHPTRAGLPAPTI

QSLNLAWNRLHAVPNLRDLPLRYLSLDGNPLAVIGPGAFAGLGGLTHLSLASLQRLPELAPS

GFRELPGLQVLDLSGNPKLNWAGAEVFSGLSSLQELDLSGTNLVPLPEALLLHLPALQSVSV

GQDVRCRRLVREGTYPRRPGSSPKVPLHCVDTRESAARGPTIL

Signal peptide:

amino acids 1-16

Transmembrane domains:

amino acids 215-232, 287-304

FIGURE 233

```
GATGGCGCAGCCACAGCTTCTGTGAGATTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGG
GGACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCA
AAACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCT
GTTCCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGC
CACGACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCAC
GGGAGGCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCC
TCTAGTCTTGCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGG
ACTGAAGACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATG
GATTTTCTGAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTA
AGGAGGACTGAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTT
GCTAAGACTCTATCTGGACAGGGTATTTAAAAACTACCAGACCCCTGACCATTATACTCTCC
GGAAGATCAGCAGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCAT
GCCCACATGACATGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCA
CTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAACTAGACATTCTTC
TGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGT
CAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACT
GTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGAT
TGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTT
TCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTT
ATTTTTTTACTTGGACATGAAACTTTAAAAAAATTCACAGATTATATTTATAACCTGACTAG
AGCAGGTGATGTATTTTTATACAGTAAAAAAAAAAAAACCTTGTAAATTCTAGAAGAGTGGCT
AGGGGGGGTTATTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGA
TATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTTGATGTGGAATT
GCACATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTAT
CTTCCAGCCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATA
CCAAAAAAAAAAAAAAAAAAAA
```

236/270

FIGURE 234

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83500
><subunit 1 of 1, 261 aa, 1 stop
><MW: 29667, pI: 8.76, NX(S/T): 0
MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRP
EIFSSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNL
QEIRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTP
DHYTLRKISSLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALG
ELDILLQWMEETE
```

Important features of the protein:
Signal peptide:
amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.
amino acids 192-195, 225-228

N-myristoylation sites.
amino acids 42-47, 46-51, 136-141

FIGURE 235

CCGTTATCGTCTTGCGCTACTGCTGA<u>ATG</u>TCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCG

CTGACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGC

CGAGCTAGCAACCTTTCCCCTGGATCTCACAAAAACTCGACTCCAAATGCAAGGAGAAGCAG

CTCTTGCTCGGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACA

GCCCTAGGGATCATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCAT

TTACAGACACGTAGTGTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTG

TGTTTGGCAAAAGTGAAGATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATG

GCTGGTGTTATTGGCCAGTTTTTAGCCAATCCAACTGACCTAGTGAAGGTTCAGATGCAAAT

GGAAGGAAAAAGGAAACTGGAAGGAAAACCATTGCGATTTCGTGGTGTACATCATGCATTTG

CAAAAATCTTAGCTGAAGGAGGAATACGAGGGCTTTGGGCAGGCTGGGTACCCAATATACAA

AGAGCAGCACTGGTGAATATGGGAGATTTAACCACTTATGATACAGTGAAACACTACTTGGT

ATTGAATACACCACTTGAGGACAATATCATGACTCACGGTTTATCAAGTTTATGTTCTGGAC

TGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCATCAAAAGCAGAATAATGAATCAACCA

CGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCGACTGACTGCTTGATTCAGGCTGT

TCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTTACCATCTTGGCTGAGAATGACCC

CTTGGTCAATGGTGTTCTGGCTTACTTATGAAAAAATCAGAGAGATGAGTGGAGTCAGTCCA

TTT<u>TAA</u>

FIGURE 236

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77568
><subunit 1 of 1, 323 aa, 1 stop
><MW: 36064, pI: 9.33, NX(S/T): 1
MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGA
RESAPYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFGKSEDEH
YPLWKSVIGGMMAGVIGQFLANPTDLVKVQMQMEGKRKLEGKPLRFRGVHHAFAKILAEGGI
RGLWAGWVPNIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTP
ADVIKSRIMNQPRDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLT
YEKIREMSGVSPF
```

Transmembrane domains:
amino acids 25-38, 130-147, 233-248

FIGURE 237

CGGACGCGTGGGCGCGGGACGCCGGCAGGGTTGTGGCGCAGCAGTCTCCTTCCTGCGCGCGC

GCCTGAAGTCGGCGTGGGCGTTTGAGGAAGCTGGGATACAGCATTTAATGAAAAATTTATGC

TTAAGAAGTAAAAATGGCAGGCTTCCTAGATAATTTTCGTTGGCCAGAATGTGAATGTATTG

ACTGGAGTGAGAGAAGAAATGCTGTGGCATCTGTTGTCGCAGGTATATTGTTTTTTACAGGC

TGGTGGATAATGATTGATGCAGCTGTGGTGTATCCTAAGCCAGAACAGTTGAACCATGCCTT

TCACACATGTGGTGTATTTTCCACATTGGCTTTCTTCATGATAAATGCTGTATCCAATGCTC

AGGTGAGAGGTGATAGCTATGAAAGCGGCTGTTTAGGAAGAACAGGTGCTCGAGTTTGGCTT

TTCATTGGTTTCATGTTGATGTTTGGGTCACTTATTGCTTCCATGTGGATTCTTTTTGGTGC

ATATGTTACCCAAAATACTGATGTTTATCCGGGACTAGCTGTGTTTTTTCAAAATGCACTTA

TATTTTTAGCACTCTGATCTACAAATTTGGAAGAACCGAAGAGCTATGGACCTGAGATCAC

TTCTTAAGTCACATTTTCCTTTTGTTATATTCTGTTTGTAGATAGGTTTTTTATCTCTCAGT

ACACATTGCCAAATGGAGTAGATTGTACATTAAATGTTTTGTTTCTTTACATTTTTATGTTC

TGAGTTTTGAAATAGTTTTATGAAATTTCTTTATTTTTCATTGCATAGACTGTTAATATGTA

TATAATACAAGACTATATGAATTGGATAATGAGTATCAGTTTTTTATTCCTGAGATTTAGAA

CTTGATCTACTCCCTGAGCCAGGGTTACATCATCTTGTCATTTTAGAAGTAACCACTCTTGT

CTCTCTGGCTGGGCACGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGG

CCGATTGCTTGAGGTCAAGTGTTTGAGACCAGCCTGGCCAACATGGCGAAACCCCATCTACT

AAAAATACAAAAATTAGCCAGGCATGGTGGTGGGTGCCTGTAATCCCAGCTACCTGGGAGGC

TGAGGCAGGAGAATCGCTTGAACCCGGGGGGCAGAGGTTGCAGTGAGCTGAGTTTGCGCCAC

TGCACTCTAGCCTGGGGGAGAAAGTGAAACTCCCTCTCAAAAAAAAGACCACTCTCAGTATC

TCTGATTTCTGAAGATGTACAAAAAAATATAGCTTCATATATCTGGAATGAGCACTGAGCCA

TAAAAGGTTTTCAGCAAGTTGTAACTTATTTTGGCCTAAAAATGAGGTTTTTTTGGTAAAGA

AAAAATATTTGTTCTTATGTATTGAAGAAGTGTACTTTTATATAATGATTTTTTAAATGCCC

AAAGGACTAGTTTGAAAGCTTCTTTTAAAAAGAATTCCTCTAATATGACTTTATGTGAGAA

FIGURE 238

MAGFLDNFRWPECECIDWSERRNAVASVVAGILFFTGWWIMIDAAVVYPKPEQLNHAFHTCG
VFSTLAFFMINAVSNAQVRGDSYESGCLGRTGARVWLFIGFMLMFGSLIASMWILFGAYVTQ
NTDVYPGLAVFFQNALIFFSTLIYKFGRTEELWT

Important features:
Signal peptide:
amino acids 1-44

Transmembrane domains:
amino acids 23-42 (type II), 60-80, 97-117, 128-148

FIGURE 239

GTTGATGGCAAACTTCCTCAAAGGAGGGGCAGAGCCTGCGCAGGGCAGGAGCAGCTGGCCCA

CTGGCGGCCCGCAACACTCCGTCTCACCCTCTGGGCCCACTGCATCTAGAGGAGGGCCGTCT

GTGAGGCCACTACCCCTCCAGCAACTGGGAGGTGGGACTGTCAGAAGCTGGCCCAGGGTGGT

GGTCAGCTGGGTCAGGGACCTACGGCACCTGCTGGACCACCTCGCCTTCTCCATCGAAGCAG

GGAAGTGGGAGCCTCGAGCCCTCGGGTGGAAGCTGACCCCAAGCCACCCTTCACCTGGACAG

G**ATG**AGAGTGTCAGGTGTGCTTCGCCTCCTGGCCCTCATCTTTGCCATAGTCACGACATGGA

TGTTTATTCGAAGCTACATGAGCTTCAGCATGAAAACCATCCGTCTGCCACGCTGGCTGGCA

GCCTCGCCCACCAAGGAGATCCAGGTTAAAAAGTACAAGTGTGGCCTCATCAAGCCCTGCCC

AGCCAACTACTTTGCGTTTAAAATCTGCAGTGGGGCCGCCAACGTCGTGGGCCCTACTATGT

GCTTTGAAGACCGCATGATCATGAGTCCTGTGAAAAACAATGTGGGCAGAGGCCTAAACATC

GCCCTGGTGAATGGAACCACGGGAGCTGTGCTGGGACAGAAGGCATTTGACATGTACTCTGG

AGATGTTATGCACCTAGTGAAATTCCTTAAAGAAATTCCGGGGGGTGCACTGGTGCTGGTGG

CCTCCTACGACGATCCAGGGACCAAAATGAACGATGAAAGCAGGAAACTCTTCTCTGACTTG

GGGAGTTCCTACGCAAAACAACTGGGCTTCCGGGACAGCTGGGTCTTCATAGGAGCCAAAGA

CCTCAGGGGTAAAAGCCCCTTTGAGCAGTTCTTAAAGAACAGCCCAGACACAAACAAATACG

AGGGATGGCCAGAGCTGCTGGAGATGGAGGGCTGCATGCCCCCGAAGCCATTT**TAG**GGTGGC

TGTGGCTCTTCCTCAGCCAGGGGCCTGAAGAAGCTCCTGCCTGACTTAGGAGTCAGAGCCCG

GCAGGGGCTGAGGAGGAGGAGCAGGGGGTGCTGCGTGGAAGGTGCTGCAGGTCCTTGCACGC

TGTGTCGCGCCTCTCCTCCTCGGAAACAGAACCCTCCCACAGCACATCCTACCCGGAAGACC

AGCCTCAGAGGGTCCTTCTGGAACCAGCTGTCTGTGGAGAGAATGGGGTGCTTTCGTCAGGG

ACTGCTGACGGCTGGTCCTGAGGAAGGACAAACTGCCCAGACTTGAGCCCAATTAAATTTTA

TTTTTGCTGGTTTTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 240

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59814
<subunit 1 of 1, 224 aa, 1 stop
<MW: 24963, pI: 9.64, NX(S/T): 1
MRVSGVLRLLALIFAIVTTWMFIRSYMSFSMKTIRLPRWLAASPTKEIQVKKYKCGLIKPCP
ANYFAFKICSGAANVVGPTMCFEDRMIMSPVKNNVGRGLNIALVNGTTGAVLGQKAFDMYSG
DVMHLVKFLKEIPGGALVLVASYDDPGTKMNDESRKLFSDLGSSYAKQLGFRDSWVFIGAKD
LRGKSPFEQFLKNSPDTNKYEGWPELLEMEGCMPPKPF
```

Important features:
Signal peptide:
amino acids 1-15

ATP/GTP-binding site motif A (P-loop).
amino acids 184-191

N-glycosylation site.
amino acids 107-110

243/270

FIGURE 241

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGAT

CCAGAAACCCATGATACCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACA

GCAAGAGAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCTC

TCACTCCTCCCTCCCTCTCTCTCTGCCTGTCCTAGTCCTCTAGTCCTCAAATTCCCAGTCCC

CTGCACCCCTTCCTGGGACACT**ATG**TTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCC

TGGCTGCAGATGGGGGTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGGCCA

GCCTCTTACCCTGAGTGTGGAAACAATGCCCAGTCGCCCATCGATATTCAGACAGACAGTGT

GACATTTGACCCTGATTTGCCTGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGC

CTTTGGACCTGCACAACAATGGCCACACAGTGCAACTCTCTCTGCCCTCTACCCTGTATCTG

GGTGGACTTCCCCGAAAATATGTAGCTGCCCAGCTCCACCTGCACTGGGGTCAGAAAGGATC

CCCAGGGGGGTCAGAACACCAGATCAACAGTGAAGCCACATTTGCAGAGCTCCA**C**ATTGTAC

ATTATGACTCTGATTCCTATGACAGCTTGAGTGAGGCTGCTGAGAGGCCTCAGGGCCTGGCT

GTCCTGGGCATCCTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTCTGAG

TCACTTGCATGAAGTCAGGCATAAAGATCAGAAGACCTCAGTGCCTCCCTTCAACCTAAGAG

AGCTGCTCCCCAAACAGCTGGGGCAGTACTTCCGCTACAATGGCTCGCTCACAACTCCCCCT

TGCTACCAGAGTGTGCTCTGGACAGTTTTTTATAGAAGGTCCCAGATTTCAATGGAACAGCT

GGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGGAGCCCTCTAAGCTTCTGGTACAGA

ACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTCATCCAAGCAGGA

TCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGCTGTCTCTG

CCTTCTCCTGGCTGTTTATTTCATTGCTAGAAAGATTCGGAAGAAGAGGCTGGAAAACCGAA

AGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCA**TAA**ATTCCTTCTCAGATACCA

TGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTG

GCCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCTTCCCCTGGACATCTCTTAGAGAG

GAATGGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTA

GGAGGAAATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAACTCTGTTTAGTTGCAGGG

GAAGTTTGGGATATACCCCAAAGTCCTCTACCCCCTCACTTTTATGGCCCTTTCCCTAGATA

TACTGCGGGATCTCTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATA

TATTTGGAAATTAAAGTTTCTGACTTT

FIGURE 242

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62812
><subunit 1 of 1, 337 aa, 1 stop
><MW: 37668, pI: 6.27, NX(S/T): 1
MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFDPDLP
ALQPHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQ
INSEATFAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRH
KDQKTSVPPFNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTL
FSTEEEPSKLLVQNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYF
IARKIRKKRLENRKSVVFTSAQATTEA
```

Important features of the protein:
Signal peptide:
amino acids 1-15

Transmembrane domain:
amino acids 291-310

N-glycosylation site.
amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins
amino acids 197-245, 104-140, 22-69

FIGURE 243

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTC

GTGGACCCAAAGGTAGCAATCTGAAAC**ATG**AGGAGTACGATTCTACTGTTTTGTCTTCTAGG

ATCAACTCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACTGGCTC

CGGATCAGGGAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTG

ATACCATTAACACAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGG

AATGACACCTGGTACCCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAAC

TGCACCCACATGTGTTACCAATTTTTGTCACACAACTTGGAGCCCAGGGCACTATCCTAAGC

TCAGAGGAATTGCCACAAATCTTCACGAGCCTCATCATCCATTCCTTGTTCCCGGGAGGCAT

CCTGCCCACCAGTCAGGCAGGGGCTAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAG

GAGCAGGTGTAAATCCTGCCACCCAGGGAACCCCAGCAGGCCGCCTCCCAACTCCCAGTGGC

ACAGATGACGACTTTGCAGTGACCACCCCTGCAGGCATCCAAAGGAGCACACATGCCATCGA

GGAAGCCACCACAGAATCAGCAAATGGAATTCAG**TAA**GCTGTTTCAAATTTTTTCAACTAAG

CTGCCTCGAATTTGGTGATACATGTGAATCTTTATCATTGATTATATTATGGAATAGATTGA

GACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATTTACCTGAAAATATTCTTGAAATT

TCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAACAATAATTCAATGGATAAAT

CTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATATTAAAACATATTTGGAAA

ACTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAA

FIGURE 244

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM

LTLGPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE

LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG

TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:
amino acids 1-16

247/270

FIGURE 245

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCG

GAGCCAGACGCTGACCACGTTCCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCG

GCAGCCGGGAGCC**ATG**CGACCCCAGGGCCCCGCCGCCTCCCCGCAGCGGCTCCGCGGCCTCC

TGCTGCTCCTGCTGCTGCAGCTGCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAG

CAAAAGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGG

GCCAGCAGGAGTGCCTGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTG

GGATCCCAGGTCGGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAG

GAGTCCTGGACACCCAACTACAAGCAGTGTTCATGGAGTTCATTGAATTATGGCATAGATCT

TGGGAAAATTGCGGAGTGTACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGT

TCAGTGGCTCACTTCGGCTAAAATGCAGAAATGCATGCTGTCAGCGTTGGTATTTCACATTC

AATGGAGCTGAATGTTCAGGACCTCTTCCCATTGAAGCTATAATTTATTTGGACCAAGGAAG

CCCTGAAATGAATTCAACAATTAATATTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAG

GAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGGTTGGCACTTGTTCAGATTACCCAAAA

GGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATCATTATTGAAGAACTACCAAAA**TA**

AATGCTTTAATTTTCATTTGCTACCTCTTTTTTTATTATGCCTTGGAATGGTTCACTTAAAT

GACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGA

CCAAAGTGTGATTTCACACTGTTTTTAAATCTAGCATTATTCATTTTGCTTCAATCAAAAGT

GGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTA

TAATTTGGAATATTGTTGTGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTAAAAAAAATA

TAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGTTAAAT

AAAAATTATTTCCAACA

FIGURE 246

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393

><subunit 1 of 1, 243 aa, 1 stop

><MW: 26266, pI: 8.43, NX(S/T): 1

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGV

PGRDGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIA

ECTFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIIYLDQGSPEMN

STINIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

2491270

FIGURE 247A

PRO   XXXXXXXXXXXXXXX   (Length = 15 amino acids)

Comparison Protein   XXXXXYYYYYYY   (Length = 12 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

5 divided by 15 = 33.3%

250/270

FIGURE 247B

PRO   XXXXXXXXXX   (Length = 10 amino acids)

Comparison Protein   XXXXXYYYYYYZZYZ   (Length = 15 amino acids)

% amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

5 divided by 10 = 50%

251/270

FIGURE 247C

PRO-DNA   NNNNNNNNNNNNNN   (Length = 14 nucleotides)

Comparison DNA   NNNNNNLLLLLLLLLL   (Length = 16 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as
determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic
acid sequence) =

6 divided by 14 = 42.9%

252/270

FIGURE 247D

PRO-DNA   NNNNNNNNNNNN   (Length = 12 nucleotides)

Comparison DNA   NNNNLLLVV   (Length = 9 nucleotides)

% nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as
determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic
acid sequence) =

4 divided by 12 = 33.3%

FIGURE 248A

```
/*
*
* C-C increased from 12 to 15
* Z is average of EQ
* B is average of ND
* match with stop is _M; stop-stop = 0; J (joker) match = 0
*/
#define  _M       -8        /* value of a match with a stop */

int     _day[26][26] = {
/*    A  B  C  D  E  F  G  H  I  J  K  L  M  N  O  P  Q  R  S  T  U  V  W  X  Y  Z */
/* A */  { 2, 0,-2, 0, 0,-4, 1,-1,-1, 0,-1,-2,-1, 0,_M, 1, 0,-2, 1, 1, 0, 0,-6, 0,-3, 0},
/* B */  { 0, 3,-4, 3, 2,-5, 0, 1,-2, 0, 0,-3,-2, 2,_M,-1, 1, 0, 0, 0,-2,-5, 0,-3, 1},
/* C */  {-2,-4,15,-5,-5,-4,-3,-3,-2, 0,-5,-6,-5,-4,_M,-3,-5,-4, 0,-2, 0,-2,-8, 0, 0,-5},
/* D */  { 0, 3,-5, 4, 3,-6, 1, 1,-2, 0, 0,-4,-3, 2,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 2},
/* E */  { 0, 2,-5, 3, 4,-5, 0, 1,-2, 0, 0,-3,-2, 1,_M,-1, 2,-1, 0, 0, 0,-2,-7, 0,-4, 3},
/* F */  {-4,-5,-4,-6,-5, 9,-5,-2, 1, 0,-5, 2, 0,-4,_M,-5,-5,-4,-3,-3, 0,-1, 0, 0, 7,-5},
/* G */  { 1, 0,-3, 1, 0,-5, 5,-2,-3, 0,-2,-4,-3, 0,_M,-1,-1,-3, 1, 0, 0,-1,-7, 0,-5, 0},
/* H */  {-1, 1,-3, 1, 1,-2,-2, 6,-2, 0, 0,-2,-2, 2,_M, 0, 3, 2,-1,-1, 0,-2,-3, 0, 0, 2},
/* I */  {-1,-2,-2,-2,-2, 1,-3,-2, 5, 0,-2, 2, 2,-2,_M,-2,-2,-2,-1, 0, 0, 4,-5, 0,-1,-2},
/* J */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */  {-1, 0,-5, 0, 0,-5,-2, 0,-2, 0, 5,-3, 0, 1,_M,-1, 1, 3, 0, 0, 0,-2,-3, 0,-4, 0},
/* L */  {-2,-3,-6,-4,-3, 2,-4,-2, 2, 0,-3, 6, 4,-3,_M,-3,-2,-3,-3,-1, 0, 2,-2, 0,-1,-2},
/* M */  {-1,-2,-5,-3,-2, 0,-3,-2, 2, 0, 0, 4, 6,-2,_M,-2,-1, 0,-2,-1, 0, 2,-4, 0,-2,-1},
/* N */  { 0, 2,-4, 2, 1,-4, 0, 2,-2, 0, 1,-3,-2, 2,_M,-1, 1, 0, 1, 0, 0,-2,-4, 0,-2, 1},
/* O */  {_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M, 0,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M,_M},
/* P */  { 1,-1,-3,-1,-1,-5,-1, 0,-2, 0,-1,-3,-2,-1,_M, 6, 0, 0, 1, 0, 0,-1,-6, 0,-5, 0},
/* Q */  { 0, 1,-5, 2, 2,-5,-1, 3,-2, 0, 1,-2,-1, 1,_M, 0, 4, 1,-1,-1, 0,-2,-5, 0,-4, 3},
/* R */  {-2, 0,-4,-1,-1,-4,-3, 2,-2, 0, 3,-3, 0, 0,_M, 0, 1, 6, 0,-1, 0,-2, 2, 0,-4, 0},
/* S */  { 1, 0, 0, 0, 0,-3, 1,-1,-1, 0, 0,-3,-2, 1,_M, 1,-1, 0, 2, 1, 0,-1,-2, 0,-3, 0},
/* T */  { 1, 0,-2, 0, 0,-3, 0,-1, 0, 0, 0,-1,-1, 0,_M, 0,-1,-1, 1, 3, 0, 0,-5, 0,-3, 0},
/* U */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */  { 0,-2,-2,-2,-2,-1,-1,-2, 4, 0,-2, 2, 2,-2,_M,-1,-2,-2,-1, 0, 0, 4,-6, 0,-2,-2},
/* W */  {-6,-5,-8,-7,-7, 0,-7,-3,-5, 0,-3,-2,-4,-4,_M,-6,-5, 2,-2,-5, 0,-6,17, 0, 0,-6},
/* X */  { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,_M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */  {-3,-3, 0,-4,-4, 7,-5, 0,-1, 0,-4,-1,-2,-2,_M,-5,-4,-4,-3,-3, 0,-2, 0, 0,10,-4},
/* Z */  { 0, 1,-5, 2, 3,-5, 0, 2,-2, 0, 0,-2,-1, 1,_M, 0, 3, 0, 0, 0, 0,-2,-6, 0,-4, 4}
};
```

FIGURE 248B

```
/*
*/
#include <stdio.h>
#include <ctype.h>

#define  MAXJMP      16       /* max jumps in a diag */
#define  MAXGAP      24       /* don't continue to penalize gaps larger than this */
#define  JMPS        1024     /* max jmps in an path */
#define  MX          4        /* save if there's at least MX-1 bases since last jmp */

#define  DMAT        3        /* value of matching bases */
#define  DMIS        0        /* penalty for mismatched bases */
#define  DINS0       8        /* penalty for a gap */
#define  DINS1       1        /* penalty per base */
#define  PINS0       8        /* penalty for a gap */
#define  PINS1       4        /* penalty per residue */

struct jmp {
        short           n[MAXJMP];      /* size of jmp (neg for dely) */
        unsigned short  x[MAXJMP];      /* base no. of jmp in seq x */
};                                      /* limits seq to 2^16 -1 */

struct diag {
        int         score;      /* score at last jmp */
        long        offset;     /* offset of prev block */
        short       ijmp;       /* current jmp index */
        struct jmp  jp;         /* list of jmps */
};

struct path {
        int     spc;            /* number of leading spaces */
        short   n[JMPS]; /* size of jmp (gap) */
        int     x[JMPS]; /* loc of jmp (last elem before gap) */
};

char            *ofile;             /* output file name */
char            *namex[2];          /* seq names: getseqs() */
char            *prog;              /* prog name for err msgs */
char            *seqx[2];           /* seqs: getseqs() */
int             dmax;               /* best diag: nw() */
int             dmax0;              /* final diag */
int             dna;                /* set if dna: main() */
int             endgaps;            /* set if penalizing end gaps */
int             gapx, gapy;         /* total gaps in seqs */
int             len0, len1;         /* seq lens */
int             ngapx, ngapy;       /* total size of gaps */
int             smax;               /* max score: nw() */
int             *xbm;               /* bitmap for matching */
long            offset;             /* current offset in jmp file */
struct  diag    *dx;                /* holds diagonals */
struct  path    pp[2];              /* holds path for seqs */

char            *calloc(), *malloc(), *index(), *strcpy();
char            *getseq(), *g_calloc();
```

255/270

FIGURE 248C

```
/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 *   where file1 and file2 are two dna or two protein sequences.
 *   The sequences can be in upper- or lower-case an may contain ambiguity
 *   Any lines beginning with ';', '>' or '<' are ignored
 *   Max file length is 65535 (limited by unsigned short x in the jmp struct)
 *   A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 *   Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"

static    _dbval[26] = {
          1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};

static    _pbval[26] = {
          1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
          128, 256, 0xFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
          1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
          1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)                                                        main
          int     ac;
          char    *av[];
{
          prog = av[0];
          if (ac != 3) {
                  fprintf(stderr,"usage: %s file1 file2\n", prog);
                  fprintf(stderr,"where file1 and file2 are two dna or two protein sequences.\n");
                  fprintf(stderr,"The sequences can be in upper- or lower-case\n");
                  fprintf(stderr,"Any lines beginning with ';' or '<' are ignored\n");
                  fprintf(stderr,"Output is in the file \"align.out\"\n");
                  exit(1);
          }
          namex[0] = av[1];
          namex[1] = av[2];
          seqx[0] = getseq(namex[0], &len0);
          seqx[1] = getseq(namex[1], &len1);
          xbm = (dna)? _dbval : _pbval;

          endgaps = 0;                    /* 1 to penalize endgaps */
          ofile = "align.out";            /* output file */

          nw();            /* fill in the matrix, get the possible jmps */
          readjmps();      /* get the actual jmps */
          print();         /* print stats, alignment */

          cleanup(0);      /* unlink any tmp files */
}
```

FIGURE 248D

```
/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */

nw()                                                                        nw
{
        char        *px, *py;          /* seqs and ptrs */
        int         *ndely, *dely;     /* keep track of dely */
        int         ndelx, delx;       /* keep track of delx */
        int         *tmp;              /* for swapping row0, row1 */
        int         mis;               /* score for each type */
        int         ins0, ins1;        /* insertion penalties */
        register    id;                /* diagonal index */
        register    ij;                /* jmp index */
        register    *col0, *col1;      /* score for curr, last row */
        register    xx, yy;            /* index into seqs */

        dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

        ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
        dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
        col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
        col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
        ins0 = (dna)? DINS0 : PINS0;
        ins1 = (dna)? DINS1 : PINS1;

        smax = -10000;
        if (endgaps) {
                for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
                        col0[yy] = dely[yy] = col0[yy-1] - ins1;
                        ndely[yy] = yy;
                }
                col0[0] = 0;        /* Waterman Bull Math Biol 84 */
        }
        else
                for (yy = 1; yy <= len1; yy++)
                        dely[yy] = -ins0;

        /* fill in match matrix
         */
        for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
                /* initialize first entry in col
                 */
                if (endgaps) {
                        if (xx == 1)
                                col1[0] = delx = -(ins0+ins1);
                        else
                                col1[0] = delx = col0[0] - ins1;
                        ndelx = xx;
                }
                else {
                        col1[0] = 0;
                        delx = -ins0;
                        ndelx = 0;
                }
```

257/276

FIGURE 248E

...nw

```
for (py = seqx[1], yy = 1; yy < = len1; py++, yy++) {
        mis = col0[yy-1];
        if (dna)
                mis += = (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
        else
                mis += = _day[*px-'A'][*py-'A'];

        /* update penalty for del in x seq;
         * favor new del over ongong del
         * ignore MAXGAP if weighting endgaps
         */
        if (endgaps || ndely[yy] < MAXGAP) {
                if (col0[yy] - ins0 > = dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else {
                        dely[yy] -= ins1;
                        ndely[yy]++;
                }
        } else {
                if (col0[yy] - (ins0+ins1) > = dely[yy]) {
                        dely[yy] = col0[yy] - (ins0+ins1);
                        ndely[yy] = 1;
                } else
                        ndely[yy]++;
        }

        /* update penalty for del in y seq;
         * favor new del over ongong del
         */
        if (endgaps || ndelx < MAXGAP) {
                if (col1[yy-1] - ins0 > = delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else {
                        delx -= ins1;
                        ndelx++;
                }
        } else {
                if (col1[yy-1] - (ins0+ins1) > = delx) {
                        delx = col1[yy-1] - (ins0+ins1);
                        ndelx = 1;
                } else
                        ndelx++;
        }

        /* pick the maximum score; we're favoring
         * mis over any del and delx over dely
         */
```

258/270

FIGURE 248F

...nw

```
id = xx - yy + len1 - 1:
if (mis > = delx && mis > = dely[yy])
        col1[yy] = mis:
else if (delx > = dely[yy]) {
        col1[yy] = delx:
        ij = dx[id].ijmp;
        if (dx[id].jp.n[0] && (!dna || (ndelx > = MAXJMP
        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
                dx[id].ijmp++;
                if (++ij > = MAXJMP) {
                        writejmps(id);
                        ij = dx[id].ijmp = 0;
                        dx[id].offset = offset;
                        offset += sizeof(struct jmp) + sizeof(offset);
                }
        }
        dx[id].jp.n[ij] = ndelx;
        dx[id].jp.x[ij] = xx:
        dx[id].score = delx:
}
else {
        col1[yy] = dely[yy];
        ij = dx[id].ijmp:

if (dx[id].jp.n[0] && (!dna || (ndely[yy] > = MAXJMP
                && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
                dx[id].ijmp++;
                if (++ij > = MAXJMP) {
                        writejmps(id);
                        ij = dx[id].ijmp = 0;
                        dx[id].offset = offset;
                        offset += sizeof(struct jmp) + sizeof(offset);
                }
        }
        dx[id].jp.n[ij] = -ndely[yy];
        dx[id].jp.x[ij] = xx:
        dx[id].score = dely[yy]:
}
if (xx = = len0 && yy < len1) {
        /* last col
        */
        if (endgaps)
                col1[yy] -= ins0+ins1*(len1-yy);
        if (col1[yy] > smax) {
                smax = col1[yy];
                dmax = id:
        }
}
}
if (endgaps && xx < len0)
        col1[yy-1] -= ins0+ins1*(len0-xx):
if (col1[yy-1] > smax) {
        smax = col1[yy-1];
        dmax = id:
}
tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);
(void) free((char *)col0);
(void) free((char *)col1);

}
```

259/270

FIGURE 248G

```
/*
 *
 * print() - only routine visible outside this module
 *
 * static:
 * getmat() - trace back best path, count matches: print()
 * pr_align() - print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() - -put a line of stars: dumpblock()
 * stripname() - strip any path and prefix from a seqname
 */

#include "nw.h"

#define SPC       3
#define P_LINE    256        /* maximum output line */
#define P_SPC     3          /* space between name or num and seq */

extern    _day[26][26];
int       olen;              /* set output line length */
FILE      *fx;               /* output file */

print()                                                                    print
{
        int     lx, ly, firstgap, lastgap;        /* overlap */

        if ((fx = fopen(ofile, "w")) == 0) {
                fprintf(stderr, "%s: can't write %s\n", prog, ofile);
                cleanup(1);
        }
        fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
        fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
        olen = 60;
        lx = len0;
        ly = len1;
        firstgap = lastgap = 0;
        if (dmax < len1 - 1) {         /* leading gap in x */
                pp[0].spc = firstgap = len1 - dmax - 1;
                ly -= pp[0].spc;
        }
        else if (dmax > len1 - 1) {    /* leading gap in y */
                pp[1].spc = firstgap = dmax - (len1 - 1);
                lx -= pp[1].spc;
        }
        if (dmax0 < len0 - 1) {        /* trailing gap in x */
                lastgap = len0 - dmax0 -1;
                lx -= lastgap;
        }
        else if (dmax0 > len0 - 1) {  /* trailing gap in y */.
                lastgap = dmax0 - (len0 - 1);
                ly -= lastgap;
        }
        getmat(lx, ly, firstgap, lastgap);
        pr_align();
}
```

260/270

FIGURE 248H

```
/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)                                    getmat
        int     lx, ly;_              /* "core" (minus endgaps) */
        int     firstgap, lastgap;    /* leading trailing overlap */
{
        int       nm, i0, i1, siz0, siz1;
        char      outx[32];
        double    pct;
        register  n0, n1;
        register char    *p0, *p1;

        /* get total matches, score
         */
        i0 = i1 = siz0 = siz1 = 0;
        p0 = seqx[0] + pp[1].spc;
        p1 = seqx[1] + pp[0].spc;
        n0 = pp[1].spc + 1;
        n1 = pp[0].spc + 1;

        nm = 0;
        while ( *p0 && *p1 ) {
                if (siz0) {
                        p1++;
                        n1++;
                        siz0--;
                }
                else if (siz1) {
                        p0++;
                        n0++;
                        siz1--;
                }
                else {
                        if (xbm[*p0-'A']&xbm[*p1-'A'])
                                nm++;
                        if (n0++ == pp[0].x[i0])
                                siz0 = pp[0].n[i0++];
                        if (n1++ == pp[1].x[i1])
                                siz1 = pp[1].n[i1++];
                        p0++;
                        p1++;
                }
        }

        /* pct homology:
         * if penalizing endgaps, base is the shorter seq
         * else, knock off overhangs and take shorter core
         */
        if (endgaps)
                lx = (len0 < len1)? len0 : len1;
        else
                lx = (lx < ly)? lx : ly;
        pct = 100.*(double)nm/(double)lx;
        fprintf(fx, "\n");
        fprintf(fx, " < %d match%s in an overlap of %d; %.2f percent similarity\n",
                nm, (nm == 1)? "" : "es", lx, pct);
```

261/270

FIGURE 248I

```
fprintf(fx, " < gaps in first sequence: %d", gapx);                                    ...getmat
if (gapx) {
        (void) sprintf(outx, " (%d %s%s)",
                ngapx, (dna)? "base":"residue", (ngapx == 1)? "":"s");
        fprintf(fx,"%s", outx);

fprintf(fx, ", gaps in second sequence: %d", gapy);
if (gapy) {
        (void) sprintf(outx, " (%d %s%s)",
                ngapy, (dna)? "base":"residue", (ngapy == 1)? "":"s");
        fprintf(fx,"%s", outx);
}
if (dna)
        fprintf(fx,
        "\n < score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
        smax, DMAT, DMIS, DINS0, DINS1);
else
        fprintf(fx,
        "\n < score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n", "
        smax, PINS0, PINS1);
if (endgaps)
        fprintf(fx,
        " < endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
        firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
        lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
else
        fprintf(fx, " < endgaps not penalized\n");
}


static          nm;              /* matches in core -- for checking */
static          lmax;            /* lengths of stripped file names */
static          ij[2];           /* jmp index for a path */
static          nc[2];           /* number at start of current line */
static          ni[2];           /* current elem number -- for gapping */
static          siz[2];
static char     *ps[2];          /* ptr to current element */
static char     *po[2];          /* ptr to next output char slot */
static char     out[2][P_LINE];  /* output line */
static char     star[P_LINE];    /* set by stars() */

/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()                                                                              pr_align
{
        int             nn;      /* char count */
        int             more;
        register        i;

        for (i = 0, lmax = 0; i < 2; i++) {
                nn = stripname(namex[i]);
                if (nn > lmax)
                        lmax = nn;

                nc[i] = 1;
                ni[i] = 1;
                siz[i] = ij[i] = 0;
                ps[i] = seqx[i];
                po[i] = out[i];
        }
```

FIGURE 248J

```
for (nn = nm = 0, more = 1; more; ) {                                    ...pr_align
        for (i = more = 0; i < 2; i++) {
                /*
                * do we have more of this sequence?
                */
                if (!*ps[i])
                        continue;

                more++;

                if (pp[i].spc) {        /* leading space */
                        *po[i]++ = ' ';
                        pp[i].spc--;
                }
                else if (siz[i]) {      /* in a gap */
                        *po[i]++ = '-';
                        siz[i]--;
                }
                else {                  /* we're putting a seq element
                                        */
                        *po[i] = *ps[i];
                        if (islower(*ps[i]))
                                *ps[i] = toupper(*ps[i]);
                        po[i]++;
                        ps[i]++;

                        /*
                        * are we at next gap for this seq?
                        */
                        if (ni[i] == pp[i].x[ij[i]]) {
                                /*
                                * we need to merge all gaps
                                * at this location
                                */
                                siz[i] = pp[i].n[ij[i]++];
                                while (ni[i] == pp[i].x[ij[i]])
                                        siz[i] += pp[i].n[ij[i]++];
                        }
                        ni[i]++;
                }
        }
        if (++nn == olen || !more && nn) {
                dumpblock();
                for (i = 0; i < 2; i++)
                        po[i] = out[i];
                nn = 0;
        }
    }
}

/*
* dump a block of lines, including numbers, stars: pr_align()
*/
static
dumpblock()                                                             dumpblock
{
        register  i;

        for (i = 0; i < 2; i++)
                *po[i]-- = '\0';
```

263 / 270

FIGURE 248K

...dumpblock

```
        (void) putc('\n', fx);
        for (i = 0; i < 2; i++) {
                if (*out[i] && (*out[i] != ' ' || *(po[i]) != '')) {
                        if (i == 0)
                                nums(i);
                        if (i == 0 && *out[1])
                                stars();
                        putline(i);
                        if (i == 0 && *out[1])
                                fprintf(fx, star);
                        if (i == 1)
                                nums(i);
                }
        }
}

/*
 * put out a number line: dumpblock()
 */
static
nums(ix)                                                                    nums
        int        ix;      /* index in out[] holding seq line */
{
        char            nline[P_LINE];
        register        i, j;
        register char   *pn, *px, *py;

        for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
                *pn = ' ';
        for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
                if (*py == ' ' || *py == '-')
                        *pn = ' ';
                else {
                        if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                                j = (i < 0)? -i : i;
                                for (px = pn; j; j /= 10, px--)
                                        *px = j%10 + '0';
                                if (i < 0)
                                        *px = '-';
                        }
                        else
                                *pn = ' ';
                        i++;
                }
        }
        *pn = '\0';
        nc[ix] = i;
        for (pn = nline; *pn; pn++)
                (void) putc(*pn, fx);
        (void) putc('\n', fx);
}

/*
 * put out a line (name, [num], seq, [num]): dumpblock()
 */
static
putline(ix)                                                                 putline
        int        ix;
{
```

264/270

FIGURE 248L

...putline

```
int              i;
register char    *px;

for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
        (void) putc(*px, fx);
for (; i < lmax+P_SPC; i++)
        (void) putc(' ', fx);

/* these count from 1:
 * ni[] is current element (from 1)
 * nc[] is number at start of current line
 */
for (px = out[ix]; *px; px++)
        (void) putc(*px&0x7F, fx);
(void) putc('\n', fx);
}



/*
 * put a line of stars (seqs always in out[0], out[1]): dumpblock()
 */
static
stars()                                                                stars
{
int              i;
register char    *p0, *p1, cx, *px;

if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
    !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
        return;
px = star;
for (i = lmax+P_SPC; i; i--)
        *px++ = ' ';

for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
        if (isalpha(*p0) && isalpha(*p1)) {

                if (xbm[*p0-'A']&xbm[*p1-'A']) {
                        cx = '*';
                        nm++;
                }
                else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
                        cx = '.';
                else
                        cx = ' ';
        }
        else
                cx = ' ';
        *px++ = cx;
}
*px++ = '\n';
*px = '\0';
}
```

265/270

FIGURE 248M

```
/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
stripname(pn)
        char    *pn;     /* file name (may be path) */
{
        register char    *px, *py;

        py = 0;
        for (px = pn; *px; px++)
                if (*px == '/')
                        py = px + 1;
        if (py)
                (void) strcpy(pn, py);
        return(strlen(pn));

}
```

stripname

266/270

FIGURE 248N

```
/*
 * cleanup() -- cleanup any tmp file
 * getseq() - read in seq, set dna, len, maxlen
 * g_calloc() - calloc() with error checkin
 * readjmps() -- get the good jmps, from tmp file if necessary
 * writejmps() - write a filled array of jmps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>


char        *jname = "/tmp/homgXXXXXX";            /* tmp file for jmps */
FILE        *fj;


int         cleanup();                              /* cleanup tmp file */
long        lseek();


/*
 * remove any tmp file if we blow
 */
cleanup(i)                                                                          cleanup
        int        i;
{
        if (fj)
                (void) unlink(jname);
        exit(i);
}


/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
char    *
getseq(file, len)                                                                   getseq
        char        *file;      /* file name */
        int         *len;       /* seq len */
{
        char                line[1024], *pseq;
        register char       *px, *py;
        int                 natgc, tlen;
        FILE                *fp;

        if ((fp = fopen(file,"r")) == 0) {
                fprintf(stderr,"%s: can't read %s\n", prog, file);
                exit(1);
        }
        tlen = natgc = 0;
        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '\n'; px++)
                        if (isupper(*px) || islower(*px))
                                tlen++;
        }
        if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
                fprintf(stderr,"%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
                exit(1);
        }
        pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
```

267/270

FIGURE 248O

```
        py = pseq + 4;                                        ...getseq
        *len = tlen;
        rewind(fp);

        while (fgets(line, 1024, fp)) {
                if (*line == ';' || *line == '<' || *line == '>')
                        continue;
                for (px = line; *px != '\n'; px++) {
                        if (isupper(*px))
                                *py++ = *px;
                        else if (islower(*px))
                                *py++ = toupper(*px);
                        if (index("ATGCU",*(py-1)))
                                natgc++;
                }
        }
        *py++ = '\0';
        *py = '\0';
        (void) fclose(fp);
        dna = natgc > (tlen/3);
        return(pseq+4);
}


char    *
g_calloc(msg, nx, sz)                                          g_calloc
        char    *msg;           /* program, calling routine */
        int     nx, sz;         /* number and size of elements */
{
        char            *px, *calloc();

        if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
                if (*msg) {
                        fprintf(stderr, "%s: g_calloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
                        exit(1);
                }
        }
        return(px);
}

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()                                                     readjmps
{
        int             fd = -1;
        int             siz, i0, i1;
        register i, j, xx;

        if (fj) {
                (void) fclose(fj);
                if ((fd = open(jname, O_RDONLY, 0)) < 0) {
                        fprintf(stderr, "%s: can't open() %s\n", prog, jname);
                        cleanup(1);
                }
        }
        for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; ; i++) {
                while (1) {
                        for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                                :
```

268/270

FIGURE 248P

...readjmps

```
if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].ijmp = MAXJMP-1;
    }
    else
        break;
}
if (i >= JMPS) {
        fprintf(stderr, "%s: too many gaps in alignment\n", prog);
        cleanup(1);
}
if (j >= 0) {
        siz = dx[dmax].jp.n[j];
        xx = dx[dmax].jp.x[j];
        dmax += siz;
        if (siz < 0) {                    /* gap in second seq */
                pp[1].n[i1] = -siz;
                xx += siz;

                /* id = xx - yy + len1 - 1
                 */
                pp[1].x[i1] = xx - dmax + len1 - 1;
                gapy++;
                ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
                siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
                i1++;
        }
        else if (siz > 0) {  /* gap in first seq */
                pp[0].n[i0] = siz;
                pp[0].x[i0] = xx;
                gapx++;
                ngapx += siz;
/* ignore MAXGAP when doing endgaps */
                siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
                i0++;
        }
    }
    else
        break;
}

/* reverse the order of jmps
 */
for (j = 0, i0--; j < i0; j++, i0--) {
        i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
        i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
        i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
        i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
        (void) close(fd);
if (fj) {
        (void) unlink(jname);
        fj = 0;
        offset = 0;
}
}
```

269/270

FIGURE 248Q

```
/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejmps(ix)                                                              writejmps
        int      ix;
{
        char     *mktemp();

        if (!fj) {
                if (mktemp(jname) < 0) {
                        fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                        cleanup(1);
                }
                if ((fj = fopen(jname, "w")) == 0) {
                        fprintf(stderr, "%s: can't write %s\n", prog, jname);
                        exit(1);
                }
        }
        (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}
```

Number	Date	Country	Number	Date	Country	Number	Date	Country
(30) 60/099,642	9 Sep/sep 1998 (09.09.1998)	US	(30) 60/101,472	23 Sep/sep 1998 (23.09.1998)	US	(30) 60/103,396	7 Oct/oct 1998 (07.10.1998)	US
(30) 60/099,741	10 Sep/sep 1998 (10.09.1998)	US	(30) 60/101,474	23 Sep/sep 1998 (23.09.1998)	US	(30) 60/103,401	7 Oct/oct 1998 (07.10.1998)	US
(30) 60/099,754	10 Sep/sep 1998 (10.09.1998)	US	(30) 60/101,475	23 Sep/sep 1998 (23.09.1998)	US	(30) 60/103,633	8 Oct/oct 1998 (08.10.1998)	US
(30) 60/099,763	10 Sep/sep 1998 (10.09.1998)	US	(30) 60/101,476	23 Sep/sep 1998 (23.09.1998)	US	(30) 60/103,678	8 Oct/oct 1998 (08.10.1998)	US
(30) 60/099,792	10 Sep/sep 1998 (10.09.1998)	US	(30) 60/101,477	23 Sep/sep 1998 (23.09.1998)	US	(30) 60/103,679	8 Oct/oct 1998 (08.10.1998)	US
(30) 60/099,808	10 Sep/sep 1998 (10.09.1998)	US	(30) 60/101,479	23 Sep/sep 1998 (23.09.1998)	US	(30) 60/103,711	8 Oct/oct 1998 (08.10.1998)	US
(30) 60/099,812	10 Sep/sep 1998 (10.09.1998)	US	(30) 60/101,738	24 Sep/sep 1998 (24.09.1998)	US	(30) 60/104,257	14 Oct/oct 1998 (14.10.1998)	US
(30) 60/099,815	10 Sep/sep 1998 (10.09.1998)	US	(30) 60/101,741	24 Sep/sep 1998 (24.09.1998)	US	(30) 60/104,987	20 Oct/oct 1998 (20.10.1998)	US
(30) 60/099,816	10 Sep/sep 1998 (10.09.1998)	US	(30) 60/101,743	24 Sep/sep 1998 (24.09.1998)	US	(30) 60/105,000	20 Oct/oct 1998 (20.10.1998)	US
(30) 60/100,385	15 Sep/sep 1998 (15.09.1998)	US	(30) 60/101,915	24 Sep/sep 1998 (24.09.1998)	US	(30) 60/105,002	20 Oct/oct 1998 (20.10.1998)	US
(30) 60/100,388	15 Sep/sep 1998 (15.09.1998)	US	(30) 60/101,916	24 Sep/sep 1998 (24.09.1998)	US	(30) 60/105,104	21 Oct/oct 1998 (21.10.1998)	US
(30) 60/100,390	15 Sep/sep 1998 (15.09.1998)	US	(30) 60/102,207	29 Sep/sep 1998 (29.09.1998)	US	(30) 60/105,169	22 Oct/oct 1998 (22.10.1998)	US
(30) 60/100,584	16 Sep/sep 1998 (16.09.1998)	US	(30) 60/102,240	29 Sep/sep 1998 (29.09.1998)	US	(30) 60/105,266	22 Oct/oct 1998 (22.10.1998)	US
(30) 60/100,627	16 Sep/sep 1998 (16.09.1998)	US	(30) 60/102,307	29 Sep/sep 1998 (29.09.1998)	US	(30) 60/105,693	26 Oct/oct 1998 (26.10.1998)	US
(30) 60/100,661	16 Sep/sep 1998 (16.09.1998)	US	(30) 60/102,330	29 Sep/sep 1998 (29.09.1998)	US	(30) 60/105,694	26 Oct/oct 1998 (26.10.1998)	US
(30) 60/100,662	16 Sep/sep 1998 (16.09.1998)	US	(30) 60/102,331	29 Sep/sep 1998 (29.09.1998)	US	(30) 60/105,807	27 Oct/oct 1998 (27.10.1998)	US
(30) 60/100,664	16 Sep/sep 1998 (16.09.1998)	US	(30) 60/102,484	30 Sep/sep 1998 (30.09.1998)	US	(30) 60/105,881	27 Oct/oct 1998 (27.10.1998)	US
(30) 60/100,683	17 Sep/sep 1998 (17.09.1998)	US	(30) 60/102,487	30 Sep/sep 1998 (30.09.1998)	US	(30) 60/105,882	27 Oct/oct 1998 (27.10.1998)	US
(30) 60/100,684	17 Sep/sep 1998 (17.09.1998)	US	(30) 60/102,570	30 Sep/sep 1998 (30.09.1998)	US	(30) 60/106,062	27 Oct/oct 1998 (27.10.1998)	US
(30) 60/100,710	17 Sep/sep 1998 (17.09.1998)	US	(30) 60/102,571	30 Sep/sep 1998 (30.09.1998)	US	(30) 60/106,023	28 Oct/oct 1998 (28.10.1998)	US
(30) 60/100,711	17 Sep/sep 1998 (17.09.1998)	US	(30) 60/102,684	1 Oct/oct 1998 (01.10.1998)	US	(30) 60/106,029	28 Oct/oct 1998 (28.10.1998)	US
(30) 60/100,919	17 Sep/sep 1998 (17.09.1998)	US	(30) 60/102,687	1 Oct/oct 1998 (01.10.1998)	US	(30) 60/106,030	28 Oct/oct 1998 (28.10.1998)	US
(30) 60/100,930	17 Sep/sep 1998 (17.09.1998)	US	(30) 60/102,965	2 Oct/oct 1998 (02.10.1998)	US	(30) 60/106,032	28 Oct/oct 1998 (28.10.1998)	US
(30) 60/100,848	18 Sep/sep 1998 (18.09.1998)	US	(30) 60/103,258	6 Oct/oct 1998 (06.10.1998)	US	(30) 60/106,033	28 Oct/oct 1998 (28.10.1998)	US
(30) 60/100,849	18 Sep/sep 1998 (18.09.1998)	US	(30) 60/103,449	6 Oct/oct 1998 (06.10.1998)	US	(30) 60/106,178	28 Oct/oct 1998 (28.10.1998)	US
(30) 60/101,014	18 Sep/sep 1998 (18.09.1998)	US	(30) 60/103,314	7 Oct/oct 1998 (07.10.1998)	US	(30) 60/106,248	29 Oct/oct 1998 (29.10.1998)	US
(30) 60/101,068	18 Sep/sep 1998 (18.09.1998)	US	(30) 60/103,315	7 Oct/oct 1998 (07.10.1998)	US	(30) 60/106,384	29 Oct/oct 1998 (29.10.1998)	US
(30) 60/101,071	18 Sep/sep 1998 (18.09.1998)	US	(30) 60/103,328	7 Oct/oct 1998 (07.10.1998)	US	(30) 60/108,500	29 Oct/oct 1998 (29.10.1998)	US
(30) 60/101,279	22 Sep/sep 1998 (22.09.1998)	US	(30) 60/103,395	7 Oct/oct 1998 (07.10.1998)	US	(30) 60/106,464	30 Oct/oct 1998 (30.10.1998)	US
(30) 60/101,471	23 Sep/sep 1998 (23.09.1998)	US				(30) 60/106,856	3 Nov/nov 1998 (03.11.1998)	US

(30) 60/106,902	3 Nov/nov 1998 (03.11.1998)	US		(30) 60/108,806	17 Nov/nov 1998 (17.11.1998)	US
(30) 60/106,905	3 Nov/nov 1998 (03.11.1998)	US		(30) 60/108,807	17 Nov/nov 1998 (17.11.1998)	US
(30) 60/106,919	3 Nov/nov 1998 (03.11.1998)	US		(30) 60/108,867	17 Nov/nov 1998 (17.11.1998)	US
(30) 60/106,932	3 Nov/nov 1998 (03.11.1998)	US		(30) 60/108,925	17 Nov/nov 1998 (17.11.1998)	US
(30) 60/106,934	3 Nov/nov 1998 (03.11.1998)	US		(30) 60/108,848	18 Nov/nov 1998 (18.11.1998)	US
(30) 60/107,783	10 Nov/nov 1998 (10.11.1998)	US		(30) 60/108,849	18 Nov/nov 1998 (18.11.1998)	US
(30) 60/108,775	17 Nov/nov 1998 (17.11.1998)	US		(30) 60/108,850	18 Nov/nov 1998 (18.11.1998)	US
(30) 60/108,779	17 Nov/nov 1998 (17.11.1998)	US		(30) 60/108,851	18 Nov/nov 1998 (18.11.1998)	US
(30) 60/108,787	17 Nov/nov 1998 (17.11.1998)	US		(30) 60/108,852	18 Nov/nov 1998 (18.11.1998)	US
(30) 60/108,788	17 Nov/nov 1998 (17.11.1998)	US		(30) 60/108,858	18 Nov/nov 1998 (18.11.1998)	US
(30) 60/108,801	17 Nov/nov 1998 (17.11.1998)	US		(30) 60/108,904	18 Nov/nov 1998 (18.11.1998)	US
(30) 60/108,802	17 Nov/nov 1998 (17.11.1998)	US				